

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/2665>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Issues of International Tax and Trade Policy
Conflict and Co-operation

by

Ben Zissimos

Thesis submitted in fullfillment of the requirements for the degree of
Doctor of Philosophy in Economics

The University of Warwick

Department of Economics

April 2003

Contents

Acknowledgements	vii
Declaration	vii
Summary	ix
1 Introduction	1
1.1 The Two Basic Questions	2
1.1.1 How can Tax Competition be Harmful?	2
1.1.2 Why Has Trade Liberalization Been (So) Gradual?	4
1.2 Outline of the arguments	5
1.2.1 Part I: A New Perspective on Tax Competition	6
1.2.2 Part II: The GATT, Regionalism, and the Postwar Trade Liberalization Process	9
1.3 Overview of The Debate on Tax Competition	16
1.3.1 ‘Tiebout’ and Leviathan Tax Competition	17
1.3.2 Capital Tax Competition	19
1.4 Overview of The Debate on Postwar Trade Liberalization	23
1.4.1 The Debate on Gradual Trade Liberalization	23
1.4.2 Regionalism versus Multilateralism	27
I A New Perspective on Tax Competition	30
2 Hotelling Tax Competition	31

CONTENTS

2

2.1	Introduction	31
2.2	The Model	39
2.3	Uniform Taxation	41
2.3.1	Stage 2: Taxes	43
2.3.2	Stage 1: Level of public good provision	48
2.4	Perfect Tax Discrimination	51
2.4.1	Stage 2: Taxes	53
2.4.2	Stage 1: Location	60
2.5	Efficiency	63
2.6	Conclusions	64
2.7	Appendix	69

II The Evolution of Postwar Trade Liberalisation 90

3 A Theory of GATT Rounds 91

3.1	Introduction	91
3.2	Optimal Tariffs, Trade Agreements and Limited Punishments	96
3.2.1	Tariffs and Welfare	96
3.2.2	Limited Punishments; Withdrawal of Equivalent Concessions	99
3.3	Symmetric Equilibrium Paths	100
3.3.1	Optimal Deviations	100
3.3.2	Efficient Equilibrium Paths and Failure to Reach Free Trade	102
3.4	Loss of Political Good Will and Gradual Tariff Reduction	106
3.4.1	Computed Paths	110
3.5	Conclusions	113
3.6	Appendix	115
3.6.1	Proof of Propositions	115
3.6.2	An Example: Quasi-linear Preferences	122

4	Why Are Trade Agreements Regional?	130
4.1	Introduction	130
4.2	A Model of Optimal Tariffs where Distance Matters	135
4.2.1	Country Location in Regions	136
4.2.2	Production and International Trade with Distance	137
4.2.3	Production-Trade Payoffs	140
4.2.4	Optimal Tariffs with Distance	141
4.2.5	Free Trade Agreements (FTAs)	142
4.3	Regions and Trade Agreements in a Non-cooperative Network . . .	144
4.3.1	FTAs as Networks	144
4.3.2	The Sponsor of an FTA	146
4.3.3	Payoffs to FTA Formation	149
4.3.4	FTA Formation with Sponsorship Costs	153
4.3.5	FTAs in Equilibrium as Nash Networks	154
4.3.6	The Dynamics of Regionalism	154
4.4	The 3-Region FTA in Equilibrium	155
4.4.1	Production-Trade Payoffs, Sponsorship Costs and Overall Payoffs in a Network Game	156
4.4.2	Equilibrium Paths; Are FTAs Stepping Blocks or Stumbling Blocks?	160
4.5	Conclusions	162
4.6	Appendix	164
5	Conclusions	177
5.1	Four Broad Themes	179
5.2	Directions for Future Research	180
	Bibliography	182

To my parents

Acknowledgements

Myrna Wooders, my main supervisor, must receive first thanks for her unstinting guidance and support, that were both a great gift and a testimony to her commitment as a social scientist. Sayantan Ghosal, Ben Lockwood and John Whalley also gave substantive supervisory support for which I am very grateful.

I also wish to thank my colleagues and friends for their advice and conversations about various parts of the thesis. Special thanks are due to Nizar Allouch, Ruth Baldry, Iwan Barankay, Paula Conconi, Amrita Dhillon, Jayasri Dutta, Isleide Maeda, Abigail McKnight, Gilluame Haeringer, Richard Higgott, Marcus Miller, Ray Riezman, Collin Rowat and John Whalley.

The biggest thanks of all go to my parents, for their unending support and for their belief, from beginning to end, that this undertaking would be worthwhile.

This research was supported financially by the Centre for the Study of Globalization and Regionalisation and the Warwick Centre for Public Economics, both at the University of Warwick. I wish to thank Richard Higgot, David Vines and Myrna Wooders for their help in securing funding.

Declaration

An earlier version of Chapter 2 was presented at a European Science Foundation Workshop held at Paris 1 in June 2002 and at a seminar at the University of Birmingham in November 2002. The chapter has now been published as Warwick Economic Research Paper 662 under joint authorship with Myrna Wooders, to acknowledge the contribution of her ideas and supervision.

An earlier version of Chapter 3 was presented at the Midwest International Economics Meetings held at University of Wisconsin in April 2001, and The WTO and World Trade II Meeting, University of Washington. The chapter has now been published as Discussion Paper 87/02 in the series published by The Centre for the

Study of Globalization and Regionalisation. As with Chapter 1, the discussion paper is published under joint authorship with Ben Lockwood, to acknowledge the contribution of his ideas and supervision.

An earlier version of Chapter 4 was presented at the Midwest International Economics Meetings held at Notre Dame University in October 2002 and was published in that form as Warwick Economic Research Paper 652.

None of the material presented in this thesis has been submitted for a degree at another university. Neither has any of the material been submitted for another degree previously.

Summary

Chapter 2, titled “Hotelling Tax Competition” shows how competition among governments for mobile firms can bring about excessive differentiation in levels of taxation and public good provision. Hotelling’s Principle of Minimum Differentiation is applied in the context of tax competition and shown to be invalid. Instead, when an equilibrium exists, differentiation of public good provision is maximized. Non-existence of equilibrium, which is possible, is a metaphor for intense tax competition. The chapter also shows that, to some extent, perfect tax discrimination presents a solution to the existence problem created by Hotelling tax competition, but that the efficiency problem of Hotelling tax competition is exacerbated.

Chapter 3 shows how the institutional rules imposed on its signatories by the GATT created a strategic incentive for countries to liberalize gradually. Free trade can never be achieved when punishment for deviation from a trade agreement is limited to a ‘withdrawal of equivalent concessions,’ the most severe form of punishment allowed (Article XXVIII). Retaliation is not allowed to entail higher tariffs than those set by the initial deviant. If, in addition, tariff bindings (Article II) limit an initial deviation from an agreement in a similar way, then efficient self-enforcing tariff reductions must proceed in a series of steps or ‘rounds’.

Chapter 4 provides an answer to the question “Why are trade agreements regional?” It argues that free trade agreements (FTAs) are regional because, in their absence, optimal tariffs are higher against (close) regional partners than (distant) countries outside the region. Optimal tariffs shift rents from foreign firms to domestic citizens. Lower transport costs imply higher rents and therefore higher tariffs. So regional FTAs have a higher payoff than non-regional FTAs. Therefore, only regional FTAs may yield positive gains when sponsoring an FTA is costly. To analyze equilibrium, standard theory of non-cooperative networks is extended to allow for asymmetric players. Naive best response dynamics show that ‘trade blocks can be stepping blocks’ for free trade.

Chapter 1

Introduction

“The best kind of economic theory has almost always reflected policy concerns, while informing policy in turn.”

(*Bhagwati, Greenaway and Panagariya* 1998)

In the areas of both international tax and trade policy, the experience of the last half a century or so has revealed that competition between governments in policy making has brought about losses in efficiency, in some areas quite significant. This dissertation takes as its starting point some worrying stylized facts about the nature of competition between national governments which existing theory does not adequately explain. It then develops aspects of existing theory in order to present ways of understanding how strategic interaction between national governments brings about losses of efficiency in ways that appear to fit the facts.

In the area of international tax policy, an acceleration in the liberalization of capital markets since 1980 appears to have brought problems of tax competition to the fore, so much so that contemporary policy debates are now focused on how to limit the degree of competition, for example through tax harmonization measures. Nevertheless, questions are still being raised as to why tax competition has been harmful, against a backdrop of conventional wisdom which suggests that competition between governments should promote efficiency.

In the area of international trade policy, by contrast, the problem of policy coordination failure was felt most acutely half a century earlier, from the 1930s until immediately after the Second World War. Since that time, trade agreements formerly under the General Agreement on Tariffs and Trade (GATT) and now under the World Trade Organization (WTO) have gone some way to resolving the coordination failure. In contrast to the field of tax competition, there is widespread agreement that unfettered trade policy intervention is harmful when it carries efficiency losses with no terms of trade (or public good) benefits to set against them. Given widespread agreement that trade policy interventions should be removed, attention in this area is focusing instead on why the post war trade liberalization process has been so gradual under the GATT. By understanding historical difficulties, it is hoped that light will be shed on why further multilateral trade liberalization is proving so difficult to achieve.

1.1 The Two Basic Questions

1.1.1 How can Tax Competition be Harmful?

In the field of tax competition, attention appears to be returning to a fundamental question: ‘In what sense can tax competition be harmful?’ It is to this question that the first part of the dissertation is addressed. According to conventional thinking based on Tiebout (1956), competition between governments is thought of as useful in that it constrains governments’ self-serving activities. This thinking applies conventional wisdom about the beneficial effects of competition between firms to the case where governments use the policy variables under their control to maximize the rents to office. Yet many would argue that tax competition between governments has done more to hinder than to enhance efficiency.

Oates (1972) argues that the result of tax competition may be a tendency to-

wards inefficiently low levels of amenity provision.¹ In an attempt to keep taxes low in order to attract business investment, government officials may hold spending below levels required for efficient amenity provision, particularly for those amenities that do not offer direct benefits to local business. This idea that amenities and taxes are driven down to inefficient levels through competition between governments has come to be known as the ‘Race to the Bottom Hypothesis’.

Yet even though capital markets have become more integrated over the last 30 years or so, recent empirical evidence has called into question the pervasiveness of the presumed worldwide race to the bottom of tax levels. For example, Baldwin and Krugman (2000) point out that tax rates have remained high in ‘the core’ of Europe in spite of far reaching market integration measures adopted by members of the European Union (EU). Capital does not appear to have flooded towards members in ‘the periphery’ despite their lower tax rates. In response, the debate over tax competition has begun to call for an explanation.

Chapter 1 of this dissertation puts forward the idea that Hotelling’s (1929) model can be adapted to understand why competition between governments in the taxation of production does not promote efficiency. In his classic article, Hotelling (1929) called into question the extent to which competition promotes efficiency when firms compete not just over prices but over product characteristics as well, and when consumers’ preferences for product characteristics vary. Chapter 1 questions, along parallel lines, the extent to which competition promotes efficiency when governments compete not just over taxes but over levels of amenity provision, and when firms’ preferences for levels of amenity provision vary. This chapter shows how competition among governments that tax the production of

¹The term ‘amenity’ (attributable in this context to unpublished research by Myrna Wooders) is used because the usual attributes of a ‘public good,’ namely non-excludability and non-rivalry, are not features of the goods that governments provide in the present analysis. Firms’ ‘preferences’ rather than firms’ technologies are referred to in order to emphasize that each firm has a clearly defined *preferred* or *ideal level of amenity provision* from which the actual level can vary.

mobile firms (the tax base in the model) can undermine efficiency.

The characterization of efficiency loss is interesting in itself. Some governments can be driven to ‘overprovide’ amenities while others ‘underprovide’ amenities in equilibrium. This helps to explain taxation in the European core and periphery for example, and possibly across the developed versus the developing world. The chapter also examines other directions in which Hotelling’s framework can be used to understand tax competition. These are explained in more detail below.

1.1.2 Why Has Trade Liberalization Been (So) Gradual?

The second general question addressed by this dissertation is ‘Why has the process of postwar trade liberalization not proceeded more rapidly towards free trade?’ Tariffs on manufactures have fallen from a trade-weighted average of about 50 percent after the war to about 5 percent today. It has required eight protracted negotiating rounds under the GATT covering the whole of the post-war period to reach current levels of openness. At the time of writing, there are significant difficulties with launching a ninth round at Doha despite a consensus that further trade liberalization would be widely beneficial, especially in agriculture and services. According to the conventional view of a trade agreement as a repeated prisoners’ dilemma, country representatives should have sat down around a table after the war and agreed upon a move more or less straight to free trade.

Meanwhile, particularly in the recent past as the multilateral liberalization process has appeared to stall, preferential trade agreements (PTAs) have proliferated. Consequently, there is a question mark over whether this proliferation of PTAs is consistent with the process of multilateral trade liberalization. Some participants in the debate hold the optimistic view that countries choose between regional and multilateral liberalization in an overall move towards free trade. Others are more pessimistic, arguing to the contrary that countries joining PTAs may be more protectionist towards non-members than they were before, leading to fragmentation

of the world trading system.

Chapter 2 of the dissertation addresses specifically the question of why trade liberalization proceeded in a series of negotiating rounds under the GATT. The GATT's institutional structure, now adopted by the WTO, is modelled as a dynamic game and the resulting (most efficient) equilibrium path is analyzed and shown to exhibit gradual trade liberalization.

Chapter 3 argues that world free trade may only be achievable via a period of regionalism. It also indicates more and less pessimistic scenarios, depending on the costs of negotiating an agreement. If the costs are relatively low then a move straight to free trade may be possible. If costs are higher then the liberalization process may stall at regionalism (or the process may not get off the ground at all if costs are very high).

1.2 Outline of the arguments

This section builds on the brief summary of questions posed and answers proposed in the previous section. While the previous section introduced the main ideas, this section discusses the development of those ideas in greater detail. The next two subsections relate quite closely to the development of the arguments in the chapters themselves. The two subsequent sections of the introduction contain overviews of the more general debates in each of the areas, putting the contributions of this dissertation into a broader context. Readers who are familiar with the literature in each of these areas may find these sections unnecessary.

The dissertation is divided in two parts. The first part motivates the need for a new perspective on tax competition and then explains the way that this is actually developed. The second part examines the dynamics of trade liberalization, both through multilateral trade liberalization under the GATT/WTO and through the formation of regional trade agreements.

1.2.1 Part I: A New Perspective on Tax Competition

The idea that politicians are self-serving, and need some form of restraint coming possibly through competition, goes all the way back to Hobbes (1651). Indeed the word *Leviathan*, the title of Hobbes' treatise, has now been adopted as the generic term for a government that uses the policy instruments at its disposal to maximize its own power. Tiebout (1956) first conjectured that competition between governments could be beneficial when he argued that efficiency is reached through competition between jurisdictional governments if citizens are able to choose, or 'vote with their feet', between jurisdictions. This idea has subsequently been established in the literature, especially by the work of Wooders (1980, 1985) and Conley and Wooders (1997, 2001); see Wooders (1999) for a comprehensive review.

A more recent preoccupation in this field is that competition between governments for mobile capital will result in a 'race to the bottom' of taxes and amenity provision. It is argued that, by taxing at a lower rate in order to prevent capital from fleeing elsewhere, each government has an incentive to engage in wasteful competition with the consequence that amenities are underprovided. Whether competition promotes or detracts from efficiency rests essentially on whether or not governments have access to efficient taxation. The weight placed on inefficient outcomes reflects recognition that taxation is not generally efficient. Wilson (1986) and Zodrow and Mieszkowski (1986) were the first to formalize the intuition of this argument, expounded by Oates (1972).

Yet empirical evidence suggests that the existing literature does not provide a comprehensive understanding of tax competition. One development is that the empirical literature remains unable to find conclusive support for the view that competition between governments promotes efficiency; see, for example, Oates (1985, 1989) and Anderson and van den Berg (1998). But the second is that recent empirical work also questions the pervasiveness of a race to the bottom in

tax rates and amenity provision; see Devereux, Griffith and Klemm (2001) for a comprehensive investigation encompassing OECD countries. More than this, there appears to be evidence that richer countries are not being forced to lower tax rates to prevent capital from migrating to poorer countries. Baldwin and Krugman (2000) present empirical evidence which suggests that countries in the European core have not needed to lower tax rates in order to prevent capital from migrating to poorer countries in the periphery. There are also suggestions that on a wider scale OECD countries have not had to lower tax rates in order to compete with developing countries. Taken together, it has been suggested that these findings call for further developments in the theory of tax competition to explain them.

As mentioned above, Chapter 1 develops a model of Hotelling tax competition to show how different governments can be driven simultaneously to ‘overprovide’ and to ‘underprovide’ amenities in equilibrium, undermining efficiency. The model may explain the patterns observed in the European core versus the periphery, and between the developed and developing worlds. Thus, the model provides a synthesized way of understanding the empirical observations outlined above.

A key element of the analysis in Chapter 1 is that firms have diverse technological requirements for levels of amenity provision. Suppose, for example, that the amenity in question is a legal system. It is generally agreed that some type of legal system will benefit a firm in its production activities and in bringing goods to market. But the ideal level of coverage differs across firms and certainly across industries. One firm’s necessary legal protection is another’s excessive red tape.²

In the previous literature, where all firms tend to have the same technological requirements for amenities, the forces of competition tend to push all governments in the same direction. With technological diversity among firms, it is not clear whether competitive forces will act similarly to push all governments in the same direction, or whether they will be pushed apart. Hotelling’s Principle of Minimum Product Differentiation predicts that governments will provide amenities at

²This idea is developed further, using other examples, in the chapter itself.

the same (inefficient) level. However, research by d'Aspremont, Gabszewicz and Thisse (1979) has called into question Hotelling's Differentiation result. Extending the intuition arising from their results on competition between firms to competition between governments suggests that competition might instead *maximize* the differentiation between governments' levels of amenity provision. Demonstrating this constitutes one of the main contributions of Chapter 1.³

An alternative possibility that arises in the framework of Chapter 1 is that an equilibrium does not exist. This 'equilibrium existence problem' is an extension of a result by d'Aspremont et al (1979) to the context of tax competition. When firms are highly responsive to a government's efforts to attract them to its jurisdiction by changing its level of amenity provision then this situation arises. Firms are more responsive to change when a move away from their ideal level of amenity provision incurs a relatively high cost. Non-existence of equilibrium in this present setting is a formal metaphor for intense tax competition. No equilibrium level of taxation exists at which governments stop undercutting each other in tax levels. Governments continually respond to each others' tax plans with successive but unending small tax reductions.

The non-existence of equilibrium characterizes quite nicely informal accounts of policy discussions amongst European policy officials, for example. They complain of continually having to look over their shoulders at the policy announcements of other governments in the EU and make counter-announcements themselves. The problem of unending tax cuts suggested by the model also seems to motivate present calls in the European policy debate for tax harmonization.

The results just discussed depend on the assumption that governments set a uniform tax schedule. That is, all firms locating in a jurisdiction must pay the same tax. The second part of Chapter 1 examines whether it is possible to resolve the problems of equilibrium existence or efficiency by allowing governments to

³See the next section for a review of the literature. See Chapter 1 for a detailed comparison of the model presented to Hotelling's original.

engage in perfect tax discrimination. As with perfect price discrimination, where firms can tailor prices to individual consumers, under perfect tax discrimination governments can tailor taxes to individual producers. Another interpretation of this alternative policy regime is that governments are able to offer tax breaks from a uniform schedule to firms in order to attract them to the jurisdiction.⁴

Bhaskar and To (2002) show that the issue of equilibrium existence in the Hotelling model is completely resolved under perfect price discrimination. In Chapter 1 it is discovered that even when governments are able to discriminate perfectly between firms in setting taxes, the equilibrium existence problem is only partially resolved. There is a larger range of values for which the cost of amenity mismatch supports an equilibrium. But even under perfect tax discrimination, if the cost of amenity mismatch is relatively high then tax competition is so intense that the system does not settle down to an equilibrium. The difference in the outcome arises because in the present analysis the level of governments' costs are endogenous, depending on the level amenity provision, whilst in the analysis of Bhaskar and To costs are exogenous as in a conventional Bertrand type framework.

1.2.2 Part II: The GATT, Regionalism, and the Postwar Trade Liberalization Process

The significant reductions in tariff rates achieved over the postwar period have been accompanied by sustained growth in international trade. World trade growth in real (volume) terms averaged 6.2 percent per year over the period 1960-1994. Whilst no single explanation is widely accepted for this growth in trade, tariff reductions achieved on a multilateral basis through the GATT are recognized as an important contributory factor. As well as generating direct welfare gains from trade, the establishment of a stable set of rules that govern the world trading sys-

⁴See Chapter 1 for a discussion of related literature on perfect price discrimination and on tax breaks.

tem is reckoned to have promoted international cooperation more widely, in areas such as FDI for example. These developments are believed to have contributed in turn to improvements in economic performance and welfare in the postwar period (Whalley and Hamilton 1996).

A Theory of GATT Rounds

These gains to trade notwithstanding, concerns have been raised over the slow pace of the liberalization process. These concerns have been amplified with an apparent slowing down of the process itself. Since the early 1980s, a literature has developed to explain why trade liberalization has been gradual. Early contributions were made from a traditional neoclassical standpoint. They tried to explain why a country would *unilaterally* (i.e. independently of behavior of other countries) wish to gradually reduce its import tariffs, based on various types of market failure within the domestic economy (see Leamer 1980 and Mussa 1986 for examples).⁵

One important aspect overlooked by all models of unilateral gradualism is the terms-of-trade motive for tariff setting. It has long been recognized (certainly since Mill 1844) that when countries have purchasing power on world markets, they can use it to improve their terms of trade using interventions such as tariffs. In such a world, countries will not adopt free trade unilaterally. Taking account of each country's own incentive to set tariffs, it is well understood that any trade agreement must be *self-enforcing*. This point was first made in the context of trade agreements by Dixit (1987). Each participating country must get a higher payoff from being in the agreement than remaining outside. And given that a country deviates from the agreement, the tariff level adopted in punishment against that country must be credible. That is, it must be in the interests of all the other parties to enact a tariff at that level.

The new literature on gradual trade liberalization plays on the credibility of

⁵See Section 1.4 for a review of the literature, where these papers and others will be discussed in greater detail.

participants' claims to adopt free trade immediately.⁶ The general idea is that initially, full liberalization cannot be self-enforcing, as the benefits of deviating from free trade are too great to be offset by any credible punishment. But if there is partial liberalization, structural economic change within the domestic economy reduces the benefits of deviation from further trade liberalization (and/or raises the costs of punishment to the deviator). This enables free trade to be approached in a series of steps.

For example, Staiger (1995) endows workers in the import competing sector with special use-it-or-loose-it skills. Initially, the government cannot credibly commit to free trade because if it reneges then it averts a contraction of the import competing sector, securing an additional payoff to the skills of workers there. These gains, together with gains in the export sector from the other country's move to free trade, make the payoff from deviation higher than from the agreement. But because the payoff to liberalization are declining at the margin, some liberalization can be committed to credibly, enabling free trade to be reached in stages.

Another way to emphasize the contrast between models of unilateral and multilateral gradualism is that in the former case a planner would liberalize gradually whilst in the latter case it would move straight to free trade; gradualism occurs due to strategic interaction between the players.

Chapter 2 asks why trade liberalization under the GATT proceeded gradually, in a series of rounds, and attributes the cause at least partly to the GATT's own institutional structure. This institutional structure is shown to create a strategic incentive for countries to proceed gradually with trade liberalization. The approach of Chapter 2 contrasts markedly with the previous literature, which concentrates on features of the domestic economy to explain gradualism. In particular, the chapter focuses on the implications for the liberalization process of the GATT rule on the *withdrawal of equivalent concessions* (WEC) as set out in Article XXVIII of the GATT charter, and on *tariff bindings* as set out in Article 2. Given that the

⁶Myopic best response tariffs are taken as the starting point.

GATT Articles have now been adopted formally by the WTO, the issues identified are likely to carry over to the new institution.⁷

Suppose that a deviant country fails to implement some agreed market access measure, whilst all other parties to the agreement proceed to do so. When the failure is discovered, under GATT rules trade partners are allowed to do no more than to withdraw market access concessions equivalent to those that the deviant failed to implement. Exactly this penalty structure is modelled in the context of a dynamic game and its implications for trade liberalization under the GATT are analyzed. In terms of the applied game theory literature, WEC imposes *partial irreversibility* on punishments in this game. This is new, in that only complete irreversibility has been analyzed in the past (Lockwood and Thomas 2002).

The first main result presented in Chapter 2 is that the WEC rule does facilitate trade liberalization but, when retaliation is limited by the WEC rule, free trade certainly cannot be reached no matter how little countries discount the future. This result contrasts markedly with conventional insights from the theory of repeated games, which indicate that free trade can be achieved, given sufficiently little discounting. The intuition behind our result is simple. A standard repeated game allows trade partners to implement the worst (credible) punishment against a deviant. In general, the WEC rule makes such severe punishments illegal. By outlawing a class of severe punishments, the WEC rule compromises efficiency. Note that for this first result, partial irreversibility is imposed only on one side of the agreement. That is to say, WEC limits only the actions of punishers.

The second main result concerns the gradualism of trade liberalization. Specifically, if punishments are constrained by the WEC rule *and* the initial deviation by any country is also constrained, then the most efficient self enforcing path of

⁷The chapter focuses on the broad sweep of postwar trade liberalization up to the end of the Uruguay round, at which point the WTO was formed. Over that period, the GATT's articles were adhered to closely by signatories. Since the WTO's inception in 1996 adherence to the rules appears to have been more limited, particularly by the EU and US.

trade liberalization is gradual; trade liberalization must take place in a series of rounds.

How is an initial deviation also constrained? Article 2 of GATT (1994) specifies that a schedule of commitments be maintained. Results of tariff negotiations are recorded as scheduled commitments in the form of tariff bindings; a permanent and irrevocable commitment that tariffs will *not rise above bound levels* for the product in question. If tariffs are raised above bound levels, then it is assumed that the deviant incurs a *loss of political good will*. Moreover, it is supposed that the loss of political good will is so costly that it is never incurred in equilibrium. This implies that the optimal deviation is simply not to cut tariffs from the previous period's level (but not to raise them either). In this situation, because punishment is limited, current tariff cuts can only be made self enforcing by the promise of future tariff reductions. Moreover, if deviation can at worst entail not raising tariffs, then it is always possible to promise liberalization over a number of future periods that would more than compensate. This is gradualism in other words. Taken together, these findings provide a way of understanding why multilateral trade liberalization proceeded in a series of rounds under the GATT, and why the process has not reached free trade.

Why Are Trade Blocks Regional; A Theory Based on Noncooperative Networks

As the process of multilateral trade liberalization has slowed down, there has been an increase in the formation of Preferential Trade Agreements (PTAs). In the period 1948-1994, GATT contracting parties notified 118 PTAs relating to trade in goods, of which 38 were notified in the five years ending in 1994. Since the completion of the Uruguay Round, 80 additional PTAs covering trade in goods and services have been notified. See Whalley and Hamilton (1996) and Sampson (1996) for more information about the recent increase in the number of preferential trade agreements. Despite the successful conclusion of the Uruguay Round, there

is still concern that increased regional integration may result in a fragmentation of the world economy into competing trade blocs.

The literature on preferential trade agreements can broadly be broken down into two phases, as suggested by Bhagwati (1991, 1993). The first phase focuses on static questions concerning the welfare effects of PTAs. This field of research was initiated by Viner (1950), who pointed out that a PTA may be ‘trade diverting’. That is, the preferential treatment of members’ goods may divert trade away from non-members that have a comparative advantage in one or more goods over both members. In a small country world, if the trade diversion effect dominates then PTA formation is efficiency reducing. Pre-Vinearian analysis of trade block formation was built on the presumption that all trade liberalization is efficiency enhancing. Current concerns about the proliferation of PTAs are founded on the possibility that trade diversion dominates.

The second phase of development in the literature focuses on whether PTAs can provide impetus to, or whether they will detract from, the worldwide freeing of trade. Bhagwati (1993) has described this as the ‘dynamic’ time-path question. In less formal terms, he asks, will trade blocks be ‘building blocks’ or ‘stumbling blocks’ in the path to free trade.

One issue that the theoretical literature has not focused much attention on is why trade agreements are almost always regional.⁸ However, as we shall see, Chapter 3 suggests that the regional nature of trade agreements may be important in understanding the dynamic time path question of regionalism.

Prominent examples of regional trade blocks are the North American Free Trade Agreement (NAFTA) and European Union (EU). In both cases, members share

⁸Only Bond (1999) considers the spacial dimension of regionalism. He compares the sustainability of multilateral versus regional trade agreements in a repeated game setting, where both types of agreement are sustained through trigger strategies. In Bond’s model, optimal tariffs are higher between closer neighbors, and this makes regional agreements easier to sustain using trigger strategies.

common borders. Wider evidence that trade blocks are predominantly regional is provided by WTO (2000), a report titled “Mapping of Regional Trade Agreements”, in which each of the 150 agreements notified to the WTO are represented in map form. Yet the theoretical literature has tended to focus on the economic implications of regional trade agreements, taking their regional nature for granted.

Some empirical investigations have sought to understand the regional nature of trade agreements. Frankel, Stein and Wei (1995) use a gravity model to show that countries behave preferentially towards close neighbors; trade volumes in the Western Hemisphere and elsewhere are greater than could be explained by ‘natural determinants’ such as distance, size and common languages. Similarly, Panagariya (1998) shows, also by taking a gravity model to the data, that transport costs alone are not sufficient to explain why trade agreements are regional.

Chapter 3 presents a theory of why there may be a strategic incentive to form regional trade blocks. In the process, the dynamic time path question is addressed by showing conditions under which trade blocks will be building blocks in the freeing of trade multilaterally and conditions when the process will stall at regionalism. The chapter argues that politicians balance the increased likelihood that they will be voted for if they coordinate a PTA against the coordination cost itself. The increased likelihood of being voted for results from conventional welfare gains to trade due to formation of the PTA. In the model, trade based gains to (close) countries of the same region are higher than gains to an agreement involving (distant) countries from different regions. The costs of bringing politicians together from different nations in order to coordinate an agreement are assumed to be proportional to the number of countries involved, and not dependent upon which countries the politicians come from. Therefore, a regional PTA may be worth coordinating whilst one involving countries from outside the region may not.

What is the basis for higher production-trade payoffs to a regional agreement? According to standard optimal tariff theory, the higher the rents made by a foreign firm in the domestic market, the more scope there is for shifting rents to domestic

citizens through the use of higher tariffs. And because trading costs increase with distance, firms make higher rents in nearby markets than those that are further away. So in the absence of an agreement, optimal tariffs are higher on imports from countries in the same region than on imports from countries of other regions. It follows that a free trade agreement (FTA) between two close neighbors brings about larger production and trade gains than between distant countries because the former entails a larger mutual tariff reduction.

Whilst standard optimal tariff theory provides a basis for individual tariff setting, a general framework is needed in which the overall structure of trade agreements in the (world) economy can be analyzed. Politicians' incentives to form international trade agreements throughout the world is formalized by adapting Bala and Goyal's (2000) model of noncooperative network formation. Bala and Goyal bring the communication networks previously modelled by others, notably Myerson (1977) and Jackson and Wolinski (1996), into a noncooperative setting.

The main result of the chapter concerns the characterization of the equilibrium FTA structure that emerges over time under different levels of sponsorship cost. Not surprisingly, if sponsorship costs are above a certain level then no FTAs will form at any point on the equilibrium path, and if they are below a certain level then world free trade will emerge straight away. It is when sponsorship costs are at an intermediate level that regionalism arises and can persist over time. Perhaps most interesting of all, a range of sponsorship costs is identified at which regionalism emerges first before free trade can be reached. In that case alone, trade blocks are indeed building blocks to the achievement of world free trade.

1.3 Overview of The Debate on Tax Competition

This section reviews the traditional literature on tax competition according to the Tiebout hypothesis and the more recent literature on capital tax competition. The section on Tiebout tax competition also reviews work on tax competition between

Leviathan governments as this can be thought of as the outcome of a departure from the perfect competition of a Tiebout model. Relevant empirical literature is also surveyed. See Wilson (1999) for an excellent comprehensive review of the wider tax competition literature. See also Berliant and Page (2001) for the state of the art on optimal income taxation and public good provision.

1.3.1 ‘Tiebout’ and Leviathan Tax Competition

As Wilson (1999) points out, Tiebout’s theory of local amenity provision also provides a theory of efficient tax competition. First the basic framework is presented, then the relationship to the idea of Leviathan governments is considered in which competition between jurisdictions is relaxed.

The Basic Framework

A jurisdiction can be thought of as a group of individuals who collectively provide amenities. To do so, the government offers local amenities that are financed through local taxes.⁹ In large economies with relatively small effective jurisdictions, there are outcomes that are efficient. That is to say, outcomes cannot be improved upon by a reorganization of individuals between jurisdictions. For example, Wooders (1985) demonstrates that when local amenities are financed by lump sum taxation and consumers can ‘opt out’ to provide the amenities for themselves, then the outcome is near-optimal, where the closeness of the outcome to efficiency depends on the costs of opting out.

There is tax competition in this model in the sense that a jurisdiction’s taxes must be kept low enough to induce individuals to reside in the jurisdiction, given the amenities that are being provided. Taxes are collected in the form of efficient lump-sum ‘head’ taxes and they are set so that each resident’s tax bill equals the cost of providing him with public amenities. This marginal-cost-pricing rule results

⁹See Wooders (1999) for a more general interpretation of the model.

in efficient migration decisions. There is now a very large literature extending these results; see Wilson (1999) and Wooders (1999) for an overview.

Tiebout tax competition models have been extended to incorporate mobile firms; see for example Richter and Wellisch (1996). There are assumed to be a large number of firms, each of which is charged a lump-sum tax that covers the cost of the amenities that it uses. The idea that firms have a large number of jurisdictions from which to choose remains important for efficiency in this setting.

Leviathan as the Converse of Efficiency under Tiebout

There is a natural presumption that if the firms in a Tiebout world were not constrained by perfect competition they would act in a self-serving monopolistic fashion. This parallels the generally expected behavior of firms as competition is relaxed. Thus, a world of Leviathan governments emerges from the Tiebout setting as a constraint is placed on the number of jurisdictions. Brennan and Buchanan (1980) argue that the size of government is excessive in the absence of competition. More recently, formal models have been constructed to show that policy-makers with Leviathan tendencies tend to behave inefficiently under limited competition; see for example Edwards and Keen (1996) and Mintz and Tulkens (1996).

The model presented in Chapter 1 is essentially a model of Leviathan governments. One way to see the formal set-up is to think of a Tiebout tax competition model with mobile firms, where firms vary in their amenity requirements, but with a limited number of jurisdictions. There are a large number of firms and in both cases taxes are effectively ‘lump sum’. It is the limit to the number of jurisdictions coupled with a variation in firms’ preferences for amenity provision that brings about a loss of efficiency under Hotelling tax competition, in contrast to the efficient outcome in the Tiebout setting.

The Empirical Literature: Looking for Leviathan

Success appears limited in the search for a relationship between aggregate government size and efficiency. Indeed, empirical tests for Leviathan by Oates (1985, 1989) and others have had difficulties in even confirming that there is a relationship between aggregate government size and the decentralization of fiscal decisions between independent governments, let alone identifying welfare implications of such a relation. More recently, Anderson and van den Berg (1998) find no evidence of a relation between fiscal decentralization and government size.

1.3.2 Capital Tax Competition

Wasteful tax competition involves some departure from the idealized setting of Tiebout models. As mentioned before, the key difference with capital tax competition is that efficient taxation is not possible. This brings about a fiscal externality. As one government changes taxes to improve the welfare of its own citizens, this lowers the welfare of citizens in other jurisdictions. The basic principle of capital tax competition is illustrated here using a simplified framework based on Wilson (1986) and Zodrow and Mieskowski (1986).

The Basic Framework

As before there are many jurisdictions. The citizens within jurisdictions are assumed to be consumers, producers and capitalists. Placing all these roles ‘under one roof’ abstracts from distributional issues. Within each jurisdiction, firms engage in competitive production, taking prices and amenities as given to produce a single output. Two factors are used in production; labor and capital. Citizens cannot move between jurisdictions, and supply their labor inelastically. Citizens also have a fixed capital endowment which they are free to rent to producers in any jurisdiction; capital is perfectly mobile. Once investment and production take place, the output is sold to residents for final consumption and to the government

as an intermediate good in production of the amenity. Consumption is financed with wages and capital income.

A nice way to illustrate the effect of capital tax competition is first to show that when labor is taxed the outcome is efficient. This is because, with labor inelastically supplied, the tax is effectively lump-sum. One justification for the assumption that efficient taxation is not feasible is that it is administratively easier to set the same tax on land and capital. More controversially, it is often argued that attempts to introduce a lump-sum tax by Mrs Thatcher, UK Prime Minister in the 1980s, met with such resistance on equity grounds that she was forced from office. When using the model to demonstrate inefficiencies through tax competition, it is assumed that tax revenue must be raised through taxation of mobile capital.

The government is assumed to be a benevolent dictator so that no loss of efficiency can be attributed to wastefulness of policy making. Even then, when mobile capital is taxed on an ad valorem or specific basis, an externality results which brings about inefficiently low taxation and amenity provision in equilibrium.

The mechanism works as follows. To finance a unit rise in spending, the government must raise taxes. As a result, the cost of capital rises, causing the demand for capital to fall. In a model where jurisdictions are assumed to be small, the interest rate remains fixed and there is no change to capital incomes. However, wages must fall to prevent firms from making negative profits. Therefore, citizens effectively pay the tax through a reduction in their wages. This tax increase must be high enough not only to pay for the increase in the marginal cost of the government spending itself but also to offset the negative effect on tax revenue of the capital outflow. Thus, citizen's income will fall as a result by more than the marginal cost of the amenity. The marginal benefit of government spending exceeds the marginal cost to compensate for the tax induced capital outflow. According to the Samuelson rule, efficiency requires that the marginal benefit of amenity provision is equal to the marginal cost. Therefore, with utility from the amenity provision declining at the margin, taxation and amenity provision must be too low.

Extensions to Basic Capital Tax Competition

The basic insights from the effects of tax competition have provoked a huge amount of interest in extensions of the basic environment and policy analysis to resolve the problem. An important extension involves analysis of situations where producers in different jurisdictions have different technologies or residents have different preferences. This gives rise to variations in tax rates. Variation in tax rates gives rise to an additional inefficiency; a fully efficient allocation cannot be achieved if tax rates differ across jurisdictions, and identical tax rates are usually not consistent with differences in amenity levels across jurisdictions arising from variations in technologies or preferences. This type of inefficiency creates a role for a central authority to put in place “corrective subsidies” which, as well as redressing interjurisdictional imbalances, can force all jurisdictions to raise taxation to the efficient level; see Wildasin (1989), DePater and Myers (1994).

Large jurisdictions have also been analyzed, as well as situations in which jurisdictions vary in size. For example, Wildasin (1988, 1989) allows changes in the demand for capital to have an effect on the interest rate. The overall effect of tax competition is the same but damped by offsetting interest rate changes. Bucovetsky (1991) and Wilson (1991) analyze asymmetric tax competition between a large jurisdiction and a small jurisdiction, distinguished by the number of residents, each with the same labor and capital endowments. They show that small jurisdictions tend to be better off under tax competition than large jurisdictions, and can even do better than in the absence of tax competition. Since the large jurisdiction demands a relatively large amount of capital, an increase in its tax rate depresses the after tax return on capital by a relatively large amount. Thus, the overall (tax inclusive) cost of capital is less sensitive to tax changes in the large jurisdiction than in the small jurisdiction. This suggests that large jurisdictions will compete less vigorously for capital through tax rate reductions and therefore end up with taxes at a higher rate.

Other situations analyzed include fiscal competition (Wildasin 1989) where governments compete not in taxes but in expenditures, producing an even more severe ‘race to the bottom’ type outcome, and situations where variations in tax policy actually creates a motive for trade (Wilson 1987).

More recently, research has identified effects to counteract the fiscal externality and bring about a ‘race to the top;’ excessive amenity provision with taxes set too high. This possibility is demonstrated by Wooders, Zissimos and Dhillon (2001) among others. Such an outcome arises in equilibrium when amenity provision enhances productivity so that a rise in taxation is associated with a rise in the demand for capital and therefore creates a positive externality between jurisdictions.

Limits to the Empirical Evidence of a ‘Race to the Bottom’

The ‘Race to the Bottom’ hypothesis has proved almost impossible to test empirically. Not only does it require the strategic interactions between jurisdictions to be tested. It also requires some estimate of the efficient level of taxation to be established before a comparison can be made to actual levels.

On a heuristic basis, much of the concern over tax competition has arisen simply because the universally quoted Statutory Tax Rate has been seen to fall across a number of countries in the last three decades or so. But Devereux, Griffith and Klemm (2001) bring together a number of different measures for the OECD countries over the period 1970-1998 to establish whether this trend is borne out when looking at other measures of capital taxation such as the Implicit Tax Rate on Corporate Profits (ITR-COR). The nature of their findings is summarized in the following quote: “The differences in the development of STAT and ITR-COR over time is striking. The former clearly fell over time while the latter did not, and if anything rose.” Mintz and Smart (2001) present and examine evidence that corporate income tax rates have remained the same or increased slightly since 1986 across provinces in Canada. Baldwin and Krugman (2000) present empirical evidence (as well as a theoretical model) which counters the idea that historically

high taxation countries have had to lower their capital tax rates across the board as European capital markets have become more integrated. Higgott (1999) draws attention to a number of other papers which cast doubt over the pervasiveness of the ‘race to the bottom’ hypothesis.

1.4 Overview of The Debate on Postwar Trade Liberalization

1.4.1 The Debate on Gradual Trade Liberalization

Since the 1980s a literature has developed to explain why world trade liberalization over the post-war period has been phased, requiring no less than eight rounds of trade talks under the GATT, spanning almost half a century. The purpose of this section is to give an overview of this literature, drawing attention to the different ways of understanding aspects of the process. The first explanations focused on market failures within the domestic economy to understand why a country might unilaterally have an incentive to liberalize gradually. The literature then moved on to take into account strategic incentives to understand why countries could not credibly commit to full liberalization immediately, but may be able to do so over time. These all focus on economic costs and benefits to liberalization that exist within the domestic economy.

Unilateral Gradualism

Early contributions were made from the traditional neoclassical standpoint. They tried to explain why a country would *unilaterally* (i.e. independently of the behavior of other countries) wish to gradually reduce its import tariffs, based on various types of market failure within the domestic economy. The first kind of explanation for unilateral gradualism is driven by the assumption that there are costs of adjustment in moving resources out of import-competing industries to other activities

(Leamer 1980, Mussa 1986). Mussa explicitly assumes convex costs of adjustment in a multi-period setting, so it follows directly that adjustment should be gradual, and the costs of adjustment are implicitly convex in Leamer.

Focusing on Mussa's explanation of unilateral gradualism, a link is drawn between the rate at which a sector contracts - due to trade liberalization - and the unemployment rate. Convexity, in this context, means that the rate of unemployment rises more than proportionally to the rate of sectoral contraction. It follows that there is an optimal gradual rate of trade liberalization. If liberalization proceeds 'too quickly', then the cost to society through unemployment is greater than the standard efficiency gains through liberalization. Mussa's approach might be criticized because unemployment in his model is not well founded in micro theory. But it is probably fair to say that there is still no general agreement on the micro-foundations of unemployment. So Mussa's starting point of simply assuming a link between sectoral contraction and unemployment, then examining the implications, has been accepted as a worthwhile and interesting contribution in this area.

Leamer's adjustment cost, measured in labor units, is proportional to the number of workers who move out of the import-competing sector. But as output is a concave function of employment, adjustment costs measured in units of output are convex i.e. 1% of the number of workers moving leads to more than a 1% decrease in output.

Unilateral gradualism can also be explained by the political economy of tariff adjustment in declining industries. Cassing and Hillman (1986) have a model where, following an exogenous negative shock in the world price, the import-competing sector can lobby the government for tariff protection. The level of the tariff is assumed to depend positively on the current level of employment in the sector. However, they focus on industry collapse (with the tariff falling to zero) rather than on gradual adjustment. Brainard and Verdier(1994) endogenize the relationship between employment and tariffs via an explicit model of lobbying and find that adjustment will be gradual (i.e. both the import tariff and employ-

ment in the declining industry fall gradually over time). However, the Brainard and Verdier model has strictly convex costs of adjustment, so a social planner would also choose gradualism. Free trade is generally consistent with theories of unilateral gradualism.

Multilateral Gradualism and the Bicycle Theory

One crucial aspect overlooked by all models of unilateral gradualism is the terms-of-trade motivation for tariff setting. It has long been recognized that when countries have purchasing power on world markets, they can use it to improve their terms of trade using trade interventions like tariffs. Only relatively recently have developments in game-theory presented trade theorists with a range of conceptual tools for thinking about the strategic interactions that result.

Taking account of each country's own incentive to set tariffs, it is well understood that any trade agreement must be *self-enforcing* (see Dixit 1987). The standard mechanism is an agreed punishment against countries that renege. This punishment must be credible. For example, if everyone knows that an optimal tariff allows at least some trade, then it would not be credible for any one country to threaten to sever all trade relations. The same incentive to deviate from no-trade exists as to deviate from free trade.

The new literature on gradual trade liberalization plays on the credibility of threatened punishments in a trade agreement, and the way that these can change as a result of the liberalization process. Different motivations have been put forward by Staiger(1995), Devereux(1997), Furusawa and Lai(1999)). The general idea is that initially, full liberalization cannot be self-enforcing, as the benefits of deviating from free trade are too great to be offset by any credible punishment. But if there is partial liberalization, structural economic change reduces the benefits of deviation from further trade liberalization (and/or raises the costs of punishment to the deviator). The individual papers differ in their description of the structural change induced by partial liberalization. As already mentioned, Staiger(1995)

endows workers in the import competing sector with specific skills, making them more productive there than elsewhere in the economy. When they move out of this sector, they lose their skills with some probability. In Devereux(1997), there is dynamic learning-by-doing in the export sector. In Furusawa and Lai(1999), there are adjustment costs incurred when labor moves between sectors. Because of the existence of adjustment costs, adjustment is not eventually to free trade in Furusawa and Lai, but to a positive tariff where the marginal world benefit from tariff reduction is equal to the resulting marginal cost of adjustment (Furusawa and Lai, Section 3). In Staiger (1995) and Devereux (1997) uninterrupted liberalization eventually results in free trade.

One idea that has been associated with gradualism is that if negotiating rounds fail then there will be a collapse back to higher levels of protectionism. This idea was first discussed informally by Bergsten (1975, page 209-24), and dubbed the ‘bicycle’ theory by Bhagwati (1988), who borrowed the term from policy circles. The issue was first addressed formally by Staiger (1995), whose model has the property that if a round of trade liberalization fails then protectionism does indeed escalate back to the level of the previous round. However, the exact nature of the factors that give rise to gradualism fundamentally affect the specific characteristics of the liberalization process. Other theories where trade liberalization is gradual do not exhibit a collapse back to higher levels of protectionism if negotiating rounds fail.

The combination of tariff-liberalization-induced resource reallocation and the ‘use-it-or-lose-it’ sector specific skills in Staiger (1995) delivers a prediction of gradualism that confirms the bicycle theory. Contrastingly, the combination of tariff liberalization induced resource reallocation and adjustment costs in Devereux (1997) and Furusawa and Lai (1999) mean that if the trade liberalization process is stopped by some unforeseen event then it is worthwhile and credible for all countries to commit to the maintenance of openness levels achieved up to that point.

Note that all of the literature outlined above focuses on aspects of the domestic economy to explain why trade liberalization is gradual. The model presented in Chapter 3, by contrast, presents a model of why trade liberalization has needed to be gradual based on a model of the GATT's institutional structure.

1.4.2 Regionalism versus Multilateralism

As pointed out above, the question of 'regionalism' has developed in two main phases.¹⁰ The first focused primarily on the welfare implications of the formation of PTAs for member and non-member countries as well as for world welfare. The second phase looked at whether regionalism is likely to promote or hinder multilateral free trade.

Most research has focused on the static effects of PTAs. These were first addressed by Viner (1950) who originated the distinction between trade diversion and trade creation. He showed that PTAs were not necessarily welfare improving, neither for member countries nor for the world as a whole. Building on the Vinerian approach, Kemp and Wan (1976) show that it is always possible to form a PTA that improves efficiency. There are two components. First, trade between members must be completely liberalized. Second, external tariffs can be adjusted to ensure that trade between members of the PTA and the rest of the world remains constant. Then the second welfare theorem can be used to show that trade between members of the PTA is efficient. And the external adjustment of tariffs eliminates trade diversion. This restores the pre-Vinerian intuition that PTA formation is welfare-improving.

However, as emphasized by Richardson (1995), Kemp and Wan's result hinges on the assumption that countries behave non-optimally. If non-member countries optimally respond to the common external tariff, members of the PTA might be worse off than in the pre-union situation.

¹⁰This literature review draws on Conconi (2002) as well as Bhagwati, Greenaway and Panagariya (1998).

Other papers have examined political incentives to form PTAs. Richardson (1993) considers the effect of an FTA in a setting where governments maximize a political support function which gives added weight to export- and import-competing producers' interests. Grossman and Helpman (1995b) use a common agency framework to analyze the way that contributions from interest groups to the government shapes the formation of FTAs.

Goyal and Joshi (2000) and Furusawa and Konishi (2002) model trade agreements as networks. Both papers show that free trade will not necessarily arise. In the case of Goyal and Joshi (2000) this is due to coordination failure. Furusawa and Konishi (2002) show that free trade fails when countries form customs unions. Both papers take a cooperative rather than a non-cooperative approach to the modelling of trade agreements as networks, and neither paper has a regional dimension. None of these previous papers considers the spacial dimension in the formation of a regional trade agreement.

Turning to the dynamic implications of regional trade block formation, Krugman (1991) suggests that the enlargement of customs unions would lead to an increase in protection against countries outside each bloc.¹¹ As a result, the world as a whole would be hurt by what appears to be a liberalizing step of promoting (preferential) free trade.¹² However, the process of regional block expansion is exogenous in Krugman's model. Welfare is simply evaluated for a division of the world into 10, 9,...2,1 blocks. To answer the so-called 'dynamic time-path question' it is necessary to examine whether forming a particular trade agreement is in the interest of the member states. Riezman (1985) was the first to analyze stable

¹¹A Customs Union is a trade agreement in which all members adopt internal free trade and coordinate the setting of a common external tariff. The EU is often cited as an example where this happens.

¹²In a monopolistically competitive framework in which provinces are divided into symmetric customs unions, Krugman (1991) shows that the division of the world into a smaller number of customs unions raises the Nash equilibrium tariff set by each block, and that world welfare is minimized when the world is divided into three symmetric blocs.

agreement structures by modelling how states choose their partners. Using the Core as a solution concept, but precluding the possibility of interstate transfers, he concludes that global free trade might not be a stable outcome. The same conclusion is reached by Kennan and Riezman (1990) and Kose and Riezman (1999). In each case, a pure exchange general equilibrium model is constructed with three countries and three goods, in which trade patterns are determined by comparative advantage considerations. Using simulation techniques to compare optimal tariffs and welfare gains in alternative agreement structures, both studies show that for certain endowment distributions customs unions can pose a threat to the multilateral trading system since, due to the improvement in their terms of trade, member countries can obtain larger welfare gains than at the free trade.

Yi (1996) uses a multi-country extension of Brander and Spencer's (1984) tariff model to describe endogenous trade block formation under imperfect competition. He addresses the issue of the sustainability of global free trade under alternative rules of Customs Union formation. Yi (1996) finds that customs unions are building blocks for global free trade if membership of a trade agreement is open to all countries, but they might be stumbling blocs if the formation of a trade bloc requires the agreement of all potential members and the number of negotiating countries exceeds a critical value. Burbidge et al (1997) describe an explicit model in which states choose their coalition partners. They show that with more than two states, incomplete federation might be the unique equilibrium, even allowing for cooperation and transfers within customs unions. All these studies are based on the assumption that trade occurs under perfect competition.

From a political economy perspective, Levy (1997) studies politicians' incentives to pursue preferential as against multilateral trade agreements in a median voter framework. Voters are affected differently by various different trade agreements depending on their endowments of labor and capital. The government picks the type of trade agreement, preferential or multilateral, that secures victory in the next election.

Part I

A New Perspective on Tax Competition

Chapter 2

Hotelling Tax Competition

2.1 Introduction

Under the conventional view of ‘government as a Leviathan’, interjurisdictional competition has come to be thought of as useful, in that it constrains governments’ self-serving activities. The view has been expounded by Brennan and Buchanan (1980), among others, who say that “... intergovernmental competition may be constitutionally ‘efficient’, regardless of the more familiar considerations of interunit spillovers examined in the orthodox theory” (p.185). This thinking applies conventional wisdom about the beneficial effects of competition between firms to the case where (Leviathan) governments behave in monopolistic fashion, using the policy variables under their control to maximize the rents to office. Yet the empirical literature remains unable to find conclusive support for this view (see, for example, Oates 1985). The problem may be that this conventional wisdom is based on a standard model, where the focus is on competition over the price of a single homogeneous good or public good. Just as firms may compete over product characteristics as well as price, governments may compete over amenities as well as taxes.

The present chapter puts forward the idea that Hotelling’s (1929) model can be

adapted to understand why competition between Leviathan governments does not promote efficiency. In his classic article, Hotelling (1929) called into question the extent to which competition promotes efficiency when firms compete not just over prices but over product characteristics as well, and when consumers' preferences for product characteristics vary. We question, along parallel lines, the extent to which competition promotes efficiency when governments compete not just over taxes but over levels of amenity provision, and when firms' preferences for levels of amenity provision vary.¹ This chapter shows how competition among governments for mobile firms can bring about excessive differentiation in levels of taxation and public good provision. Hotelling's Principle of Minimum Differentiation is applied in the context of tax competition and shown to be invalid.² Not only may there be excessive differentiation but in addition, equilibrium may fail to exist. We interpret non-existence of equilibrium as a metaphor for intense tax competition. Thus, our argument provides an explanation of why the empirical literature has remained inconclusive. This present chapter represents the first occasion on which, to my knowledge, Hotelling's model and the possible nonexistence of equilibrium have been adapted to think about amenities and taxation competition.³

¹We use the term 'amenity' because the usual attributes of a 'public good,' namely non-excludability and non-rivalry, are not features of the goods that governments provide in our analysis. We refer to firms' 'preferences' rather than firms' technologies to emphasize that each firm has a clearly defined *preferred* or *ideal level of amenity provision* from which the actual level can vary.

²In contrast, Hohaus, Konrad and Thum (1994) consider a Hotelling type model but in an analysis closer to that of Hotelling's original 1929 paper argue that competition between jurisdictions for consumers will lead to insufficient differentiation of public good quality.

³We are not the first to model interjurisdictional competition in tax and spending levels between Leviathan governments as a two stage game; this approach has been taken previously by Edwards and Keen (1996) among others. Justman, Thisse and van Ypersele (2001) treat a local public good and contrast efficiency under complete information with inefficiency under incomplete information. Their model is similar to ours in the feature that firms' preferences for public good provision are captured by their location on an interval of the real line. The Devereux, Lockwood and Redoano (2002) model also has this feature but in their model, location captures

A key element of our analysis is that firms have diverse technological requirements for levels of amenity provision. Suppose, for example, that the amenity in question is a legal system. It is generally agreed that some type of legal system will benefit a firm in its production activities and in bringing goods to market. But the ideal level of coverage differs across firms and certainly across industries. One firm's necessary legal protection is another's excessive red tape.

In broad terms, some firms within an industry operate with much less input of government provided public amenities than others. Take firms in the apparel and clothing industry as an example. Those that produce designs at the cutting edge of fashion rely more heavily on government provided amenities such as intellectual property protection, the availability of highly trained staff, and good communications networks to reach their rarefied clientele. At the other end of the spectrum are firms turning out clothing using already established patterns and brand images, for example firms producing counterfeit Levis jeans. For such firms, arguably, the more lax the levels of intellectual property protection the better. Moreover, they may have limited need of highly trained staff, and basic communications may be sufficient.

In the previous literature, where all firms tend to have the same technological requirements for amenities, the forces of competition tend to push all governments in the same direction.⁴ With technological diversity among firms, it is not clear the cost of relocation to another country rather than a "preferred level" of amenity provision that we have in our model.

⁴Situations where competition tends to push all governments away from efficiency are studied by Gordon and Wilson (1986), Wildasin (1988), Wilson (1986), Wooders, Zissimos and Dhillon (2002) and Zodrow and Miezowski (1986). In a broader context, Gordon and Wilson (1999) examine how the benefits derived by government officials from the size of the tax base can affect the design of the tax system itself. Situations where competition tends to promote efficiency are studied by Boadway, Cuff and Marceau (2002), Boadway, Pestieau and Wildasin (1989) Wildasin (1989), Wooders (1985) and Wooders, Zissimos and Dhillon (2002) among others. Oates and Schwab (1988) show that majority rule can select the efficient outcome when there is interjurisdictional competition for mobile resources. Besley and Smart (2001) argue that the

whether competitive forces will act similarly to push all governments in the same direction, or whether they will be pushed apart. Hotelling's Principle of Minimum Product Differentiation predicts that governments will provide amenities at the same (inefficient) level. However, research by d'Aspremont, Gabszwick and Thisse (1979) has called into question Hotelling's Differentiation result. Extending the intuition arising from their results on competition between firms to competition between governments suggests that competition might instead *maximize* the differentiation between governments' levels of amenity provision. Demonstrating this constitutes one of the main contributions of the present chapter.

Before considering our equilibrium analysis, we explain in a bit more detail how our model compares to Hotelling's original work. In the classic Hotelling model, consumers are located on a beach. Two ice-cream sellers chose their locations on the beach to maximize sales. Each consumer has inelastic unit demand for a *single unit* of ice cream and the only issues affecting utility are the price that the consumer has to pay for an ice-cream and the distance that he has to walk to buy it. Thus, each consumer maximizes utility by purchasing ice cream from the seller from whom the 'delivered price', including the cost of going to get the ice-cream, is the lowest.

In our model, amenity space corresponds to the beach. The further to the right that a firm is located on the interval, the higher is its preferred level of amenity provision. While Hotelling's ice cream sellers choose where to locate on the beach, in our model each government chooses a level of amenity provision in its jurisdiction. By locating within a jurisdiction, each firm is provided with the level of amenities provided by that jurisdiction. As in Hotelling's original paper, each firm is able to sell a single unit. So the only issues affecting profits

issue of whether tax competition raises or lowers efficiency depends on whether politicians are more likely to be benevolent or rent-seeking. Gordon and Wilson (2002) show that efficiency is promoted by competition when 'officials benefit by taking a smaller piece from a larger pie'. See Wilson (1999) for a comprehensive review of the earlier literature.

in our model are the tax that the firm has to pay and the difference between the firm's ideal level of amenity provision and the level actually provided in the jurisdiction where it locates. We refer to this difference between the firm's ideal level of amenity provision and the level actually provided by the government as the *degree of amenity mismatch*. The firm maximizes profits by locating in the jurisdiction where the cost of obtaining the amenity is lowest, given taxes in each jurisdiction and the degrees of amenity mismatch.

Of course, it would not be satisfactory simply to re-label Hotelling's (1929) model using the governments' variables instead of firms' variables and so on. A government's location is associated with its cost of amenity provision. In the conventional Hotelling set up, by contrast, costs of sellers are exogenous and are not linked to their location. (Applying our model to Hotelling's beach setting, it would be as if the beach gets hotter towards one end than the other, increasing a seller's costs to keep the ice cream cool.) This apparently minor modification to the set-up of Hotelling's model leads to some quite far reaching changes in its analytical properties.

The stages of the game in our model correspond to standard Hotelling analysis as well. In the first stage governments simultaneously choose the levels of amenity provision. In the second stage, after having observed each others' levels of amenity provision, governments set taxes. Of course, this ordering of events is by no means the only possible, and alternatives may well affect the outcome.⁵ As Kreps and Scheinkman (1983) argue in their study of firm behavior, the appropriateness of the set-up, or the game context as they call it, is essentially an empirical matter. Certainly, it seems reasonable to argue that governments first put in place the capacity for amenity provision in the same way that firms set up the capacity for production at the first stage. Then in the second stage they announce taxes in the

⁵In principle taxes could be set before amenity levels, both could be set at the same time, one government could behave as a Stackelberg leader at each stage and so on.

same way that firms announce prices.⁶

Aspects of our equilibrium analysis of our model carry over from d'Aspremont et al (1979) and Kreps and Scheinkman (1983). First, when equilibrium exists then, as in d'Aspremont et al, differentiation between governments in the level of amenity provision is *maximized*, contrary to the suggested prediction of Hotelling's original analysis. Given the adaptations of our model to a policy setting, however, the interpretation is different to the outcome analyzed by d'Aspremont et al. When differentiation is maximized, this implies that one government supplies no amenities at all whilst the other government supplies amenities at a maximal level.

In equilibrium governments make positive rents, as under Cournot competition, as opposed to zero rents, as under Bertrand competition. The result is particularly striking for the jurisdiction that supplies no amenities at all even though it levies a positive tax. This arises as a result of the monopolistic power that each government has over location within its jurisdiction. Each firm must have a jurisdictional location in order to produce, and the government of that jurisdiction is able to exploit its resultant power when setting taxes.

Recent research has drawn attention to the persistent differences between what have come to be known as the core and the periphery of Europe. The core includes Benelux, France, Germany and Italy. The periphery includes Spain, Portugal, Ireland and Greece. For example, Baldwin and Krugman (2000) show how significant differences in taxes, and therefore amenity provision, have persisted over the last thirty five years or so, even as capital markets have become more integrated.⁷ Stylistically, the core of Europe could be associated with the high tax high

⁶In a wider setting, beyond the context of our model, governments have the power to tax citizens first and then spend the revenue on public services. But multinational firms can be thought of as more like customers, choosing to locate in a jurisdiction only once the amenity is available for use there.

⁷The theoretical model presented by Baldwin and Krugman (2000) motivates persistent differences in taxation and amenity provision between the core and periphery by allowing the core to move first in the policy setting game. First mover advantage gives them an incentive to act as

amenity providing government of our model and the periphery could be associated with the low tax low amenity providing government. Our equilibrium prediction that differentiation between levels of amenity provision is maximized provides a way of understanding why these observed differences between the European core and periphery have persisted.

To fix ideas, return to the example of the clothing and apparel industry. Our analysis may suggest that the forces of competition drive governments in the European core to over-provide amenities in order to attract (or retain) the companies of haute couture, that have a preference for a relatively high level of amenity provision. Given that a government in the European periphery provides amenities at a relatively low level (none at all in this stylized setting) and sets taxes relatively low, a government in the core cannot do any better by mimicking the periphery government. At the same time, the amenities offered by core governments are not sufficiently important to the production technologies of more standard clothing producers, and it is not worth paying the higher taxes of the core in order to be able to locate there.

As noted above, it is a possibility in our framework that an equilibrium does not exist. When firms are highly responsive to a government's efforts to attract them to its jurisdiction by changing its level of amenity provision then this situation arises. Firms are more responsive to change when a move away from their ideal level of amenity provision incurs a relatively high cost. Non-existence of equilibrium in this present setting is a formal metaphor for intense tax competition. No equilibrium level of taxation exists at which governments stop undercutting each other in tax levels.⁸

Stackelberg leaders, setting high taxes and providing a high level of amenities. This then creates agglomeration of industry through external economies, as highlighted in the recent literature on the economics of location; see Krugman (1998) for a review.

⁸At first sight, this appears to imply that rents fall to or below zero. This is not the case. As shown by d'Aspremont et al for prices, no equilibrium exists when a small reduction in taxes is sufficient to attract all firms to the jurisdiction. In this case, governments keep responding to

In light of the equilibrium existence issue raised by the foregoing analysis, perfect tax discrimination is analyzed to examine the extent to which it provides a solution. As with perfect price discrimination, where firms can tailor prices to individual consumers, under perfect tax discrimination governments can tailor taxes to individual producers. One interpretation is that governments are able to offer tax breaks from a uniform schedule to firms in order to attract them to the jurisdiction.⁹ Bhaskar and To (2002) show that the issue of equilibrium existence in the Hotelling model is completely resolved under perfect price discrimination. In our model we find that allowing governments to discriminate perfectly in setting taxes only partially resolves the equilibrium existence problem. There is a larger range of values for which the cost of amenity mismatch supports an equilibrium. But even under perfect tax discrimination, if the cost of amenity mismatch is relatively high then tax competition is so intense that the system does not settle down to an equilibrium.

Finally, under conditions where equilibrium exists, efficiency implications of the respective regimes are compared. The same inefficiency exists under Hotelling tax/amenity competition with uniform taxes as under the conventional Hotelling model analyzed by d'Aspremont, Gabszwick and Thisse (1979). Product differentiation is maximal and therefore excessive. Research by Spence (1976) (in the context of firms) suggests that giving governments more power to discriminate between firms in terms of the taxes they are charged will increase and possibly maximize efficiency. Bhaskar and To (2002) show that this reasoning carries over to the original Hotelling framework of firm location and production. But we find that for our model efficiency loss is *worse* under perfect tax discrimination. In equilibrium, both governments offer no amenities at all. This exerts a high effi-

each other's tax plans with smaller and smaller but unending tax reductions.

⁹Earlier research by Bond and Samuelson (1986), Black and Hoyt (1989), Haaparanta (1996) and King, McAfee and Welling (1993) model situations where governments offer some firms more favorable treatment than others but they either model competition for a single firm or assume firms' technological requirements for amenities are identical.

ciency loss on firms that have a high public good requirement, and leads to a lower aggregate level of efficiency. There is a key difference in Bhaskar and To's analysis of firms. In their setting, each firm has the same fixed level of cost. In our analysis, recall that governments' costs depend on their level of amenity provision. Under perfect tax discrimination, the higher-amenity-providing government loses out to the lower one because of the higher cost of provision. This creates a unilateral incentive to deviate from any relatively high level of amenity provision, bringing about a 'race to the bottom' of taxes and amenity provision.

The chapter proceeds as follows. Section 2 sets out the basic model. Sections 3, and 4 examine Hotelling tax/amenity competition, looking for existence of subgame perfect equilibrium under uniform taxation and perfect tax discrimination respectively. Section 5 then compares the welfare implications of the regimes when equilibrium exists. Section 6 concludes.

2.2 The Model

We adapt Hotelling's model to the problem of tax competition. The governments of two countries, A and B , compete over taxes and the level of amenity provision in attempting to persuade firms to locate in their jurisdictions. These governments are assumed to be Leviathans, maximizing the rents to office through amenity provision. There is a continuum of firms uniformly distributed on a (non-empty) interval $s \in [0, z]$.¹⁰ The position (fixed in technology space) of each firm in the interval $s \in [0, z]$ reflects its ideal level of amenity provision to facilitate production.

The location on the interval $[0, z]$ of the two governments A and B is given by variables a and b respectively. The variable a measures the distance from 0 and b measures the distance from z ; $a + b \leq z$, $a \geq 0$, $b \geq 0$. The location of the government determines the level of amenity provision to each firm in the

¹⁰This could be generalized so that there are multiple firms at each point on the interval, but this would not add insight.

jurisdiction; a to each firm in Jurisdiction A and $(z - b)$ to each firm in Jurisdiction B . The tax on the firm positioned at s is τ_{As} if the firm locates in Jurisdiction A and τ_{Bs} if it locates in Jurisdiction B .

In conventional Hotelling fashion, each firm is able to sell a single unit and to charge price $p = d$. The cost function for the firm at $s \in [0, z]$ is given by

$$c_s = \begin{cases} c + \tau_{As} + k|s - a| & \text{if the firm locates in Jurisdiction } A \\ c + \tau_{Bs} + k|s - (z - b)| & \text{if the firm locates in Jurisdiction } B. \end{cases}$$

If the firm at s locates in A , for example, it must pay private cost c , and tax τ_{As} . The firm's position s indicates its *ideal level of amenity provision*. The *degree of amenity mismatch* of the firm positioned at s is given by the distance of the firm from the location of the government. For example, if the firm locates in A then the degree of amenity mismatch is given by $|s - a|$. The impact on costs of a divergence from this ideal level of amenity provision would then be captured by the term $k|s - a|$, where k parameterizes the impact of the degree of amenity mismatch on costs. We refer to k as the *cost of amenity mismatch* for short. Firm profits are given by $\pi_s \equiv p - c_s$. To focus the analysis on location decisions, it will be assumed throughout that p is high enough to ensure that all firms make positive profits.

The model described above is illustrated in Figure 1. The figure shows the set of firms $s \in [0, z]$. The locations of governments A and B at points a and b are also pictured. The point \hat{s} shows the position of the marginal firm choosing to locate in Jurisdiction A . The firm at \hat{s} is indifferent between Jurisdiction A and B because it makes the same profits in either.

To summarize, in terms of their technological requirements for amenity provision, firms' positions are fixed, but firms are able to pick their preferred jurisdiction to maximize profits. Each government, on the other hand, is able to pick its level of amenity provision but obviously its jurisdiction (A or B) is fixed.

2.3 Uniform Taxation

Under a *uniform tax game*, each government is able only to set a uniform tax on the firms that choose to locate in its jurisdiction. Government A sets a tax $\tau_{As} = \tau_A$ and makes rents of $\tau_A - a$ on each firm in its jurisdiction while Government B sets a tax $\tau_{Bs} = \tau_B$ and makes rents of $\tau_B - (z - b)$ on each firm in its jurisdiction. It is a condition of equilibrium that $\tau_A - a \geq 0$. The same condition applies to Government B ; $\tau_B - (z - b) \geq 0$.

Given that a and b measure the distances of governments A and B from 0 and z respectively, and that $a + b \leq z$, it must be the case that $a < \hat{s} < b$. Then

$$-\tau_A - k|s - a| = -\tau_B - k|s - (z - b)|$$

Hence

$$\hat{s}(\tau_A, \tau_B) = \frac{\tau_B - \tau_A}{2k} + \frac{(z - b + a)}{2}.$$

A firm may be closer to one government, say Government A , in terms of its degree of amenity mismatch; $|s - a| < |s - (z - b)|$. But if the net cost of public good procurement is sufficiently low, the firm may choose to locate in Jurisdiction B , accepting a higher degree of amenity mismatch; formally, this holds when $-\tau_B - k|s - (z - b)| < -\tau_A - k|s - a|$. Thus if it could set $\tau_B < \tau_A$ by a sufficiently wide margin, Government B could attract any firm $s \in [0, z]$.

The solution to the governments' problems, the levels of amenity provision and the taxes that they set, can now be determined in the outcome of a game. The two governments, A and B , play respective pure strategies $\tau_A \in \mathbb{R}_+$ and $\tau_B \in \mathbb{R}_+$.¹¹

¹¹It will be assumed throughout that mixed strategies in tax rates are not available to governments. This is generally deemed to be an acceptable assumption in the applied literature on policy setting in a perfect information environment. Intuitively, it would not be regarded as reasonable for a government to announce a policy of randomizing over tax rates. Admittedly, there may be more complex tax setting environments in which mixed strategies would make more sense. Developments in that direction are left for further research.

Payoffs are given by the ‘rents to office’ which are defined by the following rent functions:

$$r_A(\tau_A, \tau_B) = \begin{cases} z(\tau_A - a) & \text{if } \tau_A < \tau_B - k(z - a - b) \\ \frac{1}{2}(z + a - b)(\tau_A - a) - \frac{1}{2k}(\tau_A - a)\tau_A + \frac{1}{2k}(\tau_A - a)\tau_B & \text{if } |\tau_A - \tau_B| \leq k(z - a - b) \\ 0 & \text{if } \tau_A > \tau_B + k(z - a - b) \end{cases}$$

$$r_B(\tau_A, \tau_B) = \begin{cases} z(\tau_B - (z - b)) & \text{if } \tau_B < \tau_A - k(z - a - b) \\ \frac{1}{2}(z - a + b)(\tau_B - (z - b)) - \frac{1}{2k}(\tau_B - (z - b)) + \frac{1}{2k}(\tau_B - (z - b))\tau_A & \text{if } |\tau_A - \tau_B| \leq k(z - a - b) \\ 0 & \text{if } \tau_B > \tau_A + k(z - a - b) \end{cases}.$$

If $\tau_A < \tau_B - k(z - a - b)$ then Government A attracts all firms to locate in Jurisdiction A and it makes overall rents of $z(\tau_A - a)$; see the first line on the right hand side of the rent function $r_A(\tau_A, \tau_B)$. If Government A sets $\tau_A > \tau_B + k(z - a - b)$ then no firm finds it profitable to locate in Jurisdiction A and there are no rents to be made from office there; see the last line on the right hand side of $r_A(\tau_A, \tau_B)$. Over the *firm sharing interval*, $|\tau_A - \tau_B| \leq k(z - a - b)$, some firms locate in each of the jurisdictions. Then rents for Government A are given by $r_A(\tau_A, \tau_B) = (\tau_A - a)\hat{s}$, the reduced form of which is given in the middle line on the right hand side of $r_A(\tau_A, \tau_B)$.

The ‘rent function’ of Government A is shown in Figure 2 for a fixed value $\bar{\tau}_B$. It shows two discontinuities, which occur at the taxes $\tau_A = \tau_B - k(z - a - b)$ and $\tau_A = \tau_B + k(z - a - b)$. At each discontinuity, all firms are indifferent between locating in either of the two jurisdictions. This property of the pay-off function, that it has two discontinuities, is familiar from the previous literature on stability in Hotelling’s model (see d’Aspremont, Gabszwick, and Thisse 1979, for example).

It is clear that $r_A(\tau_A, \tau_B)$ is linear in τ_A for $\tau_A < \tau_B - k(z - a - b)$ and equal to zero for $\tau_A > \tau_B + k(z - a - b)$. To see that $r_A(\tau_A, \tau_B)$ is strictly concave over the firm sharing interval, note that $\partial^2 r_A(\tau_A, \tau_B) / \partial \tau_A^2 = -1/k$ over the interval $|\tau_A - \tau_B| \leq k(z - a - b)$. The same holds for $r_B(\tau_A, \tau_B)$.

Amenity provision and tax setting is modelled as a two stage game. In the first stage, the governments A and B simultaneously determine their levels of amenity provision. In the second stage, they set taxes. Once the governments' decisions have been taken, firms take taxes and amenities as given and choose their geographical locations (ie, A or B) to maximize profits. Each of the two stages constitutes a subgame for which it is possible to determine whether or not there exists a Nash equilibrium. Then we say that there exists a *subgame-perfect Nash equilibrium* if the players' strategies constitute a Nash equilibrium in every subgame. It follows that if in either period there exists no Nash equilibrium in pure strategies then there is no subgame perfect Nash equilibrium (in pure strategies). We identify conditions on the existence of a subgame perfect Nash equilibrium of this game.

2.3.1 Stage 2: Taxes

The purpose of this section is to solve for Stage 2, where the location of the two governments is taken as fixed at distances a and b from the ends of the interval $[0, z]$ (ie at distances a from 0 and b from z respectively). As we shall see, when a and b are 'too close' an equilibrium fails to exist.

For given locations a and b , a strategy τ_A^* of Government A is a *best response tax* against a strategy τ_B when it maximizes $r_A(\tau_A, \tau_B)$ on the whole of \mathbb{R}_+ . A *Nash equilibrium in taxes* is a pair (τ_A^*, τ_B^*) for which (i) τ_A^* is a best response to τ_B^* and vice-versa (ii) $\tau_A^* \geq a$ and $\tau_B^* \geq z - b$.

By standard results, if the rent functions were everywhere continuous and concave, then existence of a unique best response would be guaranteed. Because

the rent function for each government is discontinuous, the usual first and second order conditions cannot be used to find best responses. However, it will be possible to show that when a Nash equilibrium does exist it is unique. Moreover, the tax choice of each jurisdiction maximizes its rents, and maximal rents are given by the maximum of the rent function on the firm sharing interval $|\tau_A - \tau_B| \leq k(z - a - b)$; see Figure 2.

The first step is to solve for the tax that maximizes rent on the firm sharing interval.

Lemma 1. *Assume governments play a uniform tax game. For given τ_B , the unique tax that maximizes $r_A(\tau_A, \tau_B)$ on the firm sharing interval is*

$$\tau_A(\tau_B; a, b, k, z) = k \left(\frac{a + \tau_B}{2k} + \frac{(z + a - b)}{2} \right).$$

For given τ_A , the unique tax τ_B that maximizes $r_B(\tau_A, \tau_B)$ on the firm sharing interval is

$$\tau_B(\tau_A; a, b, k, z) = k \left(\frac{(z - b) + \tau_A}{2k} + \frac{(z - a + b)}{2} \right).$$

If $\tau_A(\tau_B; a, b, k, z)$ and $\tau_B(\tau_A; a, b, k, z)$ are set simultaneously, then they can be solved for simultaneously to obtain:

$$\tau_A(a, b, k, z) = \frac{1}{3} (2a + (z - b) + (a - b)k + 3kz);$$

$$\tau_B(a, b, k, z) = \frac{1}{3} (2(z - b) + a + (b - a)k + 3kz).$$

As the rent function is strictly concave on the firm sharing interval, each government has a unique maximizing tax on that interval, taking the tax set by the other government as given. From the positive sign that the tax of the other government takes on the right hand side, it is clear that taxes are strategic complements.

The second part of the result says that when both governments set $\tau_A(\tau_B; a, b, k, z)$ and $\tau_B(\tau_A; a, b, k, z)$ simultaneously, each can be expressed strictly in terms of model parameters; $\tau_A(a, b, k, z)$ and $\tau_B(a, b, k, z)$. Of course, if this is the case

then $\tau_A(\tau_B; a, b, k, z)$ and $\tau_B(\tau_A; a, b, k, z)$ are mutual best responses and constitute a Nash equilibrium point. This will only be the case, though, if, given the other government's tax, there is no tax outside the firm sharing interval that yields higher rent.

It is straightforward to check whether the highest payoff is yielded by the rent maximizing tax on the firm sharing interval or some other tax that attracts all firms to the jurisdiction. This check is performed in the next result.

Lemma 2. *Under a uniform tax game, the tax $\tau_A(\tau_B; a, b, k, z)$ that maximizes $r_A(\tau_A, \tau_B)$ on the firm sharing interval $|\tau_A - \tau_B| \leq k(z - a - b)$ is a best response to τ_B if and only if, for any τ_B and $\varepsilon > 0$,*

$$r_A(\tau_A(\tau_B; a, b, k, z), \tau_B) \geq z(\tau_B - k(z - a - b) - a - \varepsilon).$$

Similarly, the tax $\tau_B(\tau_A; a, b, k, z)$ that maximizes $r_B(\tau_A, \tau_B)$ on the firm sharing interval $|\tau_A - \tau_B| \leq k(z - a - b)$ is a best response to τ_A if and only if, for any τ_A and $\varepsilon > 0$,

$$r_B(\tau_A, \tau_B(\tau_A; a, b, k, z)) \geq z(\tau_A - k(z - a - b) - (z - b) - \varepsilon).$$

The only meaningful alternative to a best response tax in the firm sharing interval is a best response tax that attracts all firms to the jurisdiction.¹² In the first inequality, $r_A(\tau_A(\tau_B; a, b, k, z), \tau_B)$ gives the maximum rent for Jurisdiction A on the firm sharing interval, and $z(\tau_B - k(z - a - b) - a - \varepsilon)$ gives the rent from setting a tax low enough to attract all firms to A. In the case of Government A, for example, this tax is $\tau_A = \tau_B - k(z - a - b) - \varepsilon$. The second inequality gives a parallel expression for Jurisdiction B. Recall that a firm would accept a higher degree of amenity mismatch if the tax were low enough to make the net cost of public good procurement lower. At the tax implied by the right hand side

¹²From Lemma 1, $\tau_A(\tau_B; a, b, k, z)$ and $\tau_B(\tau_A; a, b, k, z)$ are both non-negative. So given that each country has a positive share of firms rents cannot be negative, and raising taxes to the point where no firms are attracted to the jurisdiction can be rejected as a possible best response.

of the inequality all firms, even those which have a smaller degree of amenity mismatch with Government B , would locate in Jurisdiction A because of the more favorable tax. Lemma 2 says that the $\tau_A(\tau_B; a, b, k, z)$ that maximizes rents on the firm sharing interval is a best response tax if and only if no tax $\tau_A = \tau_B - k(z - a - b) - \varepsilon$ exists that yields higher rents.

We are now ready to state conditions on the existence and uniqueness of a Nash equilibrium in the second stage, taking locations a and b , and parameters k and z as given. It will show that an equilibrium of this Stage 2 subgame exists if and only if each government has a best response tax that is on its firm sharing interval.

Proposition 1. *Assume governments play a uniform tax game, and that a and b are fixed on the interval $[0, z]$, with $a + b \leq z$, $a \geq 0$, $b \geq 0$. For $a + b = z$, both governments are at the same location and there exists an equilibrium in which $\tau_A^* = a$, $\tau_B^* = z - b$.*

For $a + b < z$ there exists an equilibrium point if and only if the two following conditions hold:

$$(C1): r_A(\tau_A^*(\tau_B^*; a, b, k, z), \tau_B^*) \geq z(\tau_B^* - k(z - a - b) - a - \varepsilon) \Leftrightarrow \\ \frac{((a - b)k + (z - a - b) + 3kz)^2}{18k} \geq \frac{z(2(a + 2b)k + 2(z - a - b) - 3\varepsilon)}{3}$$

$$(C2): r_B(\tau_B^*(\tau_A^*; a, b, k, z), \tau_A^*) \geq z(\tau_A^* - k(z - a - b) - (z - b) - \varepsilon) \Leftrightarrow \\ \frac{((b - a)k - (z - a - b) + 3kz)^2}{18k} \geq \frac{z(2(2a + b)k - 2(z - a - b) - 3\varepsilon)}{3}$$

Whenever it exists, an equilibrium point is determined uniquely by the taxes

$$\tau_A^*(a, b; k, z) = \frac{1}{3}(2a + (z - b) + (a - b)k + 3kz); \\ \tau_B^*(a, b; k, z) = \frac{1}{3}(2(z - b) + a + (b - a)k + 3kz).$$

The first line of conditions C1 and C2 is familiar from Lemma 2. Here in Propo-

sition 1, however, equilibrium values have been substituted. The Proposition establishes conditions under which the taxes that maximize rents in the firm sharing intervals of each government are mutual best responses. It also shows that if such taxes are not mutual best responses then equilibrium fails to exist.

The second line of C1 and C2 gives conditions for existence and uniqueness in terms of model parameters a , b , k and z . As stated, these reduced form conditions are not transparent. However, in the next section where stage 1 of the game is solved it will become clear that $a = 0$ and $b = 0$ are the only candidates for equilibrium. Checking that C1 and C2 hold having made these substitutions for a and b is straightforward.

The intuition behind Proposition 1 can be understood as follows. First, the situation where $a + b = z$ is directly analogous to a standard model of Bertrand competition, where each government offers the same amenity level. So there exists a Bertrand equilibrium, which is efficient in that neither government makes rents.

Second, in the situation where $a + b < z$, so that governments supply differing levels of amenities, Proposition 1 says that an equilibrium exists if and only if the tax set by each government is in the firm sharing interval. Suppose not. Suppose at the rent maximizing tax, where firms are shared, one government can do better by setting a tax sufficiently low to attract all firms to its jurisdiction. Then the other government has an incentive to undercut the first. The undercutting process continues ad infinitum and equilibrium is never reached. This does not mean that taxes become infinitely negative. The budget surplus condition always holds. As d'Aspremont, Gabszwick and Thisse (1979) show for firms, only a small tax reduction is needed in such a situation to attract all firms to the local jurisdiction.

Although the basic insight of d'Aspremont et al (1979) carries over the present context of tax competition, the analysis in the present context is more complicated. The additional complications arise because our model allows governments to differ by offering different levels of amenities. The choice of amenity level affects the government's cost of provision. Recall that this is somewhat different from the

conventional Hotelling set-up where firms offer a product that is homogeneous in all respects other than the location at which it is supplied. Varying location does not affect a firm's costs in Hotelling's conventional model. In our setting, by contrast, varying location does affect a government's cost of amenity provision. This adds an extra part to the process of solving for equilibrium. Lemma 1 shows that taxes become strategic complements in the firm sharing interval. That is, τ_B enters positively in $\tau_A(\tau_B; a, b, k, z)$ and τ_A enters positively in $\tau_B(\tau_A; a, b, k, z)$. This is different from the analysis of d'Aspremont et al, where there is no strategic substitution or complementarity at all.

Because taxes are strategic complements in the firm sharing interval, conditions C1 and C2 are somewhat less transparent than in d'Aspremont et al (1979). A nice feature of their formative analysis is that each condition is shown to depend in a clear way on the difference between a and b . When a and b are 'too close' equilibrium fails to exist. It is through this route that d'Aspremont et al (1979) introduce their main result; that Hotelling's Principle of Minimum Differentiation fails to hold. Contrastingly, the relationship between a and b in C1 and C2 cannot be discerned so clearly in the present analysis. However, a nice clear alternative demonstration of the present model's failure to exhibit the Principle of Minimum Differentiation will be given in the next section.

2.3.2 Stage 1: Level of public good provision

We now solve for Stage 1, defining an equilibrium in locations, which determines the level of public good provision by the respective governments. For Government A , the rent function is $r_A(\tau_A, \tau_B)$. Using the equilibrium values $\tau_A^* = \tau_A^*(a, b; k, z)$ and $\tau_B^* = \tau_B^*(a, b; k, z)$ that we derived for Stage 2, the rent function for Government A can be written as follows:

$$r_A(\tau_A^*(a, b; k, z), \tau_B^*(a, b; k, z)) = r_A(a, b; k, z).$$

Similarly, the rent function for Government B can be written as follows:

$$r_B(\tau_A^*(a, b; k, z), \tau_B^*(a, b; k, z)) = r_B(a, b; k, z).$$

A location a^* of Government A is a *best response* against a location b when it maximizes $r_A(a, b; k, z)$ on the whole of \mathbb{R}_+ . A *Nash equilibrium in locations* is a pair (a^*, b^*) such that a^* is a best response against b^* and vice-versa.

Substituting $\tau_A^* = \frac{1}{3}(2a + (z - b) + (a - b)k + 3kz)$ and $\tau_B^* = \frac{1}{3}(2(z - b) + a + (b - a)k + 3kz)$ into $r_A(\tau_A^*, \tau_B^*) = (\tau_A^* - a)\hat{s}(\tau_A^*, \tau_B^*)$, Government A 's problem in Stage 1 of the game can be written as follows:

$$\max_a r_A(a, b; k, z) = \frac{((a - b)k + (z - a - b) + 3kz)^2}{18k}.$$

Similarly, Government B 's problem can be written

$$\max_b r_B(a, b; k, z) = \frac{((b - a)k - (z - a - b) + 3kz)^2}{18k}.$$

The game played between these two governments has an unconventional but nonetheless appealing form. To demonstrate that the Principle of Maximum Differentiation holds, we will first show that the second derivative of the rent function is everywhere *nonnegative*. This implies that, when the first derivative of the rent function is strictly negative, each government's rents will be maximized by moving as far from the location of the other government (in amenity provision space) as possible.

Lemma 3 shows how the second order condition of the government's problem in the first stage is non-negative.

Lemma 3. *Assume a uniform tax game.*

$$\frac{\partial^2 r_A(a, b; k, z)}{\partial a^2} = \frac{(k - 1)^2}{9k}, \quad \frac{\partial^2 r_B(a, b; k, z)}{\partial b^2} = \frac{(k + 1)^2}{9k}$$

Lemma 3, along with (C1) and (C2), are used to check that in equilibrium rents to office cannot be increased by changing location.

Our next proposition shows that the maximal differentiation result of d'Aspremont, Gabszewicz, and Thisse (1979) extends, when equilibrium exists, to the present model.

Proposition 2. *There exists a unique subgame perfect Nash equilibrium in pure strategies of a uniform tax game if and only if $0 < k \leq \frac{1}{7}$. If such an equilibrium exists then it is characterized (uniquely) by the point $a^* = b^* = 0$.*

This result shows that an equilibrium exists only if and only if the costs of amenity mismatch are relatively low ($k \leq 1/7$). If an equilibrium exists then differentiation in amenity provision is maximized. (Recall that a measures the distance from 0 and b measures the distance from z .) To see why it is the case, consider the incentives to deviate from the equilibrium $a^* = b^* = 0$. As governments move away from each other they increase the degree of differentiation of the amenity level that they offer. This in turn softens the degree of tax competition that they face, which increases the rents that can be made from any given level of amenity provision. If the costs of amenity mismatch are relatively high ($k > \frac{1}{7}$) then more firms switch to the government that is closer to the centre of the interval, producing a unilateral incentive to deviate from $a = b = 0$. However, if governments have an incentive to deviate from $a = b = 0$ then equilibrium fails to exist. The reason is that as the governments move closer to the centre of the interval, tax competition becomes more intense. That is, the incentive for one government to reduce taxes and in so doing attract all firms to its jurisdiction increases. No equilibrium level exists at which taxes stop falling. Thus, in non-existence of equilibrium we have a formal metaphor for intense tax competition.¹³

Comparing the results obtained here with those of d'Aspremont, Gabszewicz and Thisse (1979), in their earlier analysis, when mismatch costs were linear, a subgame perfect Nash equilibrium failed to exist for all parameter values. D'Aspremont et

¹³As mentioned in the introduction, this does not mean that taxes become infinitely negative. The budget surplus condition always holds. As d'Aspremont et al (1979) show for firms, only a small tax reduction is needed in such a situation to attract all firms to the local jurisdiction.

al were able to demonstrate existence of equilibrium only in an alternative model where mismatch costs were quadratic. In our present model with just a linear framework, we have been able to show that existence of equilibrium or otherwise depends on the cost parameter associated with mismatch k . Quadratic costs are not required to show existence. This difference of model properties arises out of the differences of our model to the standard Hotelling set-up. In our model location affects rents directly through costs. For example, for Jurisdiction A , $r_A(\tau_A, \tau_B) = (\tau_A - a) \hat{s}(\tau_A, \tau_B)$. The analogous expression in the conventional Hotelling set-up would be $r_A(\tau_A, \tau_B) = \tau_A \hat{s}(\tau_A, \tau_B)$. The differences in model behavior are driven by the feature that location affects rents directly through costs.

Given the adaptations of the Hotelling model to our policy context, the interpretation is different to that provided by d'Aspremont et al (1979) as well. In the conventional model, other than location there is no difference between the characteristics of the products being supplied by the two firms. When differentiation is maximized this simply means that the goods are supplied at different locations. Here in the context of this present chapter, when differentiation is maximized this implies that one government supplies no amenities at all whilst the other government supplies amenities at a maximal level.

2.4 Perfect Tax Discrimination

In a *perfect tax discrimination game*, each government is able to set an individualized tax for each firm $s \in [0, z]$. Each government is able to set an individual tax for the firm at s , in the same way as firms that perfectly price discriminate are able to set an individualized price for each consumer. Unlike in the previous section where each government set a single tax which all firms locating in that jurisdiction had to pay, now each government is able to set a different tax for each firm. The two governments A and B then engage in Bertrand competition

separately for each firm. In this section, we consider the extent to which perfect tax discrimination resolves the problems of existence of equilibrium under uniform taxation.

Thinking more loosely, there is an alternative interpretation of the perfect tax discrimination game. If there existed a uniform tax schedule in each country then this model of perfect tax discrimination could be seen as capturing the incentive for governments to offer individualized tax breaks to firms in order to attract them to the jurisdiction.

For each firm $s \in [0, z]$, the two governments, A and B , play respective strategies $\tau_{As} \in \mathbb{R}_+$ and $\tau_{Bs} \in \mathbb{R}_+$. The rent functions to competition for this single firm are given as follows:

$$r_{As}(\tau_{As}, \tau_{Bs}) = \begin{cases} (\tau_{As} - a) & \text{if } \tau_{As} < \tau_{Bs} + k(|(z - b) - s| - |s - a|) \\ 0 & \text{if } \tau_{As} > \tau_{Bs} + k(|(z - b) - s| - |s - a|) \end{cases}.$$

$$r_{Bs}(\tau_{As}, \tau_{Bs}) = \begin{cases} (\tau_{Bs} - (z - b)) & \text{if } \tau_{Bs} < \tau_{As} + k(|s - a| - |(z - b) - s|) \\ 0 & \text{if } \tau_{Bs} > \tau_{As} + k(|s - a| - |(z - b) - s|) \end{cases}.$$

The rent received by each government when $\tau_{As} - \tau_{Bs} = k(|(z - b) - s| - |s - a|)$ will be specified presently.

Each of the rent functions has a single discontinuity. An example of $r_{As}(\tau_{As}, \tau_{Bs})$ is shown in Figure 3. For any $\tau_{As} < \tau_{Bs} + k(|(z - b) - s| - |s - a|)$, the firm finds it profitable to locate in Jurisdiction A . That is, the difference between the costs of amenity mismatch $k(|(z - b) - s| - |s - a|)$ across the two jurisdictions is more than offset by the difference in the taxes. The government makes rent $\tau_{As} - a$ on the firm at s . If $\tau_{As} > \tau_{Bs} + k(|(z - b) - s| - |s - a|)$, the difference in taxes more than offsets the difference between the costs of amenity mismatch across the jurisdictions, and the firm locates in Jurisdiction B . Then, obviously, the government

makes rents of zero on the firm at s .

The firm is just indifferent between the two jurisdictions at the point $\tau_{As} = \tau_{Bs} + k(|(z - b) - s| - |s - a|)$. This is the point of discontinuity in $r_{As}(\tau_{As}, \tau_{Bs})$ shown in Figure 3. The difference in the costs of amenity mismatch and the difference in the taxes across the two jurisdictions is exactly equal. We need to specify how firm s will decide its location when it is just indifferent between jurisdictions. The following assumption stipulates that either jurisdiction is chosen with probability one half.

A1: If $\tau_{As} - \tau_{Bs} = k(|(z - b) - s| - |s - a|)$ for $s \in [0, z]$ then s is indifferent between A and B and chooses each jurisdiction with probability $\frac{1}{2}$. The expected rent for Government A is $\frac{1}{2}(\tau_{As} - a)$ and the expected rent for Government B is $\frac{1}{2}(\tau_{Bs} - (z - b))$.

Again, as in Section 3, the level of amenity provision and tax setting is modelled as a two stage game. As before, the governments A and B simultaneously determine their levels of amenity provision in Stage 1, and set taxes in Stage 2. Each of the two periods constitutes a subgame for which it is possible to determine whether there exists a Nash equilibrium. Then there exists a *subgame-perfect Nash equilibrium* if the governments' strategies constitute a Nash equilibrium in every subgame. As in the previous section, it follows that if in either period there exists no Nash equilibrium then there is no subgame perfect Nash equilibrium.

2.4.1 Stage 2: Taxes

As usual, Stage 2 is solved for first, where the location of the two governments is taken as fixed at distances a and b from the ends of the interval $[0, z]$. For given locations a and b and for a given firm $s \in [0, z]$, a strategy τ_{As}^* of Government A is a *best response* against a strategy τ_{Bs} when it maximizes $r_{As}(\tau_{As}, \tau_{Bs})$ on \mathbb{R}_+ . A *Nash equilibrium in taxes for firm s* is a pair $(\tau_{As}^*, \tau_{Bs}^*)$ for which (i) τ_{As}^* is a best response to τ_{Bs}^* and vice-versa. (ii) $\tau_A^* \geq a$ and $\tau_B^* \geq z - b$.

Let $T_A = \{\tau_{As}\}_{s \in [0, z]}$ be a *tax schedule* for Government A , consisting of one tax for each firm, and similarly let $T_B = \{\tau_{Bs}\}_{s \in [0, z]}$ be a tax schedule for Government B . A pair of tax schedules, T_A^* and T_B^* is a *Nash equilibrium in taxes* if for each $s \in [0, z]$ the pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium in taxes for firm s .

The literature on entry deterrence through pricing strategy has had to broach the issue of what constitutes a best response when payoff functions defined by the game are discontinuous and do not have a well defined maximum (in the sense that first derivatives are not equal to zero). This issue carries over to the present context where the payoff function is increasing up to the discontinuity; see Figure 3. In a model of continuous strategy choices, such a payoff function does not have a well defined maximum because, for any strategy chosen by a player, there is always a strategy that yields a slightly higher payoff. Consider, for example, the present setting where any choice of ε implies a tax $\tau_{As} = \tau_{Bs} + k(|(z - b) - s| - |s - a|) - \varepsilon > 0$, ($\varepsilon > 0$) and rent $r_{As} = \tau_{As} - a$. Government A could choose a smaller value for ε (whilst still maintaining $\varepsilon > 0$) thereby setting a higher tax and earning higher rent.

Dasgupta and Maskin (1986) provide a way of resolving this issue by defining (discrete) strategy choices over a grid. In such a framework, ε has a smallest value defined by the distance between grid lines. Their approach has gained substantive support in the literature and, in the present setting, has intuitive appeal. Let $\varepsilon > 0$ be thought of as the smallest monetary unit; one cent in the Euro zone or the US and a penny in Canada or the UK, for example. With a smallest money unit, the minimum amount by which one government can undercut the other is well defined as ε . Then $r_{As}(\tau_{As}, \tau_{Bs})$ has a well defined maximum. Strategies can be made continuous by making the distance between grid lines arbitrarily small.¹⁴

¹⁴A formal game theoretic treatment, along the lines of Dasgupta and Maskin (1986), could be developed for Hotelling Tax Competition. In such an approach, discrete taxes would be defined over a grid, with distance between grid lines equal to ε , and ε would then be allowed to become arbitrarily small. Inclusion of such a derivation would not contribute substantively to the results

For our purposes, we simply define a ‘limit tax’ for a firm s as a tax very close to but less than the tax that would make the firm indifferent between the two jurisdictions. To formalize a limit tax, let $\varepsilon > 0$ be given. For a particular firm s , a tax τ_{Bs} , and amenity levels a and b satisfying $z - b > a$, the *limit tax for Government A*, τ_{As}^{lim} , is given by:

$$\tau_{As}^{\text{lim}} = \tau_{Bs} + k(|(z - b) - s| - |s - a|) - \varepsilon.$$

Analogously, for a particular firm s , a tax τ_{As} , and amenity levels a and b satisfying $z - b > a$, the *limit tax for Government B*, τ_{Bs}^{lim} , is given by:

$$\tau_{Bs}^{\text{lim}} = \tau_{As} + k(|s - a| - |(z - b) - s|) - \varepsilon.$$

Notice that the limit tax is not relevant for the case $z - b = a$, where competition between governments is analogous to Bertrand competition in homogeneous products. When setting a limit tax in Stage 2, Government A effectively takes a , b , k , s , z and τ_{Bs} , as given, so we write the limit tax τ_{As}^{lim} as a function of ε only; $\tau_{As}^{\text{lim}}(\varepsilon)$. Analogously, for the limit tax of Government B we write $\tau_{Bs}^{\text{lim}}(\varepsilon)$.

The notion of limit tax that we introduce here extends to a tax policy setting the idea of a limit price originally introduced by Bain (1956). Bain suggested that pricing strategies could be used to discourage entry.¹⁵ Bhaskar and To (2002) show that pricing strategies can be used to discourage entry into a market that is defined geographically. A particular firm can supply its nearby market relatively cheaply because it can provide the good in question at relatively low delivery cost. Then the limit price is the highest price the firm can charge without making it possible that we discuss in the present paper. Such a formal treatment of limit pricing by firms has been undertaken by Chowdhury (2002). The price that maximizes the payoff as the grid size becomes small is defined as the limit price.

¹⁵Spence (1977) re-interprets limit pricing as competition in capacities, where an incumbent accumulates a large capacity and thus charges a low price, deterring entry. Milgrom and Roberts (1982) formulate a model based on informational asymmetry, where an incumbent charges a low price to signal that profits in the market are low.

for other more distant firms to profitably supply the market. For limit pricing to be a best response, profits must be maximized if the firm is the local market's sole supplier.

In the policy setting of this present chapter, tax strategies can be used to discourage competition for a particular set of firms defined not in terms of their location but in terms of their degree of amenity mismatch. A particular government can provide an amenity to a firm with a relatively small degree of amenity mismatch at a tax that enables the firm to make relatively high profits; the closer is the level of amenity provision to the firm's ideal the higher are the profits that the firm makes, all else equal. From the point of view of one government, the limit tax is the highest tax that it can set for a firm while making it impossible for the other government to profitably provide an amenity on more favorable terms. The limit tax then maximizes the rent that can be made.

Using the definitions of limit taxes, we can now characterize the best response for each government in Stage 2.

Lemma 4. *Consider a perfect tax discrimination game and assume A1 holds. Fix a and b so that $z - b > a$.*

If, for some firm $s \in [0, z]$, $a < \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then for $\varepsilon > 0$ sufficiently small Government A's unique best response is $\tau_{As}^ = \tau_{As}^{\lim}(\varepsilon)$. If $a \geq \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then $\tau_{As}^* = a$ is a best response for Government A.*

If, for some firm $s \in [0, z]$, $z - b < \tau_{As} + k(|s - a| - |(z - b) - s|)$ then for $\varepsilon > 0$ sufficiently small Government B's unique best response is $\tau_{Bs}^ = \tau_{Bs}^{\lim}(\varepsilon)$. If $z - b \geq \tau_{As} + k(|s - a| - |(z - b) - s|)$ then $\tau_{Bs}^* = z - b$ is a best response for Government B.*

The first part of the result says that if, from Government A's point of view, the degree of amenity mismatch with a firm at s is small relative to that firm's mismatch with Government B, then it is a best response for Government A to set a limit tax for that firm. Formally, if $a < \tau_{Bs} + k(|(z - b) - s| - |s - a|)$

then $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$. Notice that $\tau_{Bs} + k(|(z - b) - s| - |s - a|)$ is decreasing in the degree of amenity mismatch $|s - a|$, making the condition more likely to hold if s is close to a . For given tax and location of Government B , Government A limit taxes the firm so it just prefers to locate in A . If, on the other hand, $a \geq \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then Government A can do no better than to set $\tau_{As}^* = a$. Clearly, setting $\tau_{As}^* < a$ would make negative rents. And given that the firm is not attracted to A at $\tau_{As}^* = a$, then it certainly will not find $\tau_{As}^* > a$ more attractive. The second part of the result states that parallel arguments hold for the best response of Government B .

In Lemma 4 and in the following, we mean by ‘ $\varepsilon > 0$ sufficiently small’ that the smallest monetary unit is small enough to enable the government that has the smaller degree of amenity mismatch with a given firm to undercut the other government using taxes. That is, we rule out the possibility that one government is closer in amenity space to a firm than the other government but not able to undercut the other on taxes and still make positive rents because the smallest monetary unit is too large. The formal bound on the size of ε is established in the proof.

The best responses determined above are now used to define equilibrium in the next two propositions.

Proposition 3. *Consider Stage 2 of a perfect tax discrimination game, with a and b fixed on the interval $[0, z]$. Assume A1 holds and that $a + b \leq z$, $a \geq 0$, $b \geq 0$. If $k < 1$ then for $\varepsilon > 0$ sufficiently small there exists a unique Nash equilibrium in taxes for this stage of the perfect tax discrimination game. A unique Nash equilibrium in taxes for each firm $s \in [0, z]$ is determined by the following taxes:*

if $a + b = z$,

$$\tau_{As}^* = \tau_{Bs}^* = a = z - b;$$

if $a + b < z$,

$$\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon), \tau_{Bs}^* = z - b.$$

Proposition 3 can be explained as follows. If $a + b = z$ then we have the standard Bertrand case. If $a + b < z$ then, with relatively low costs of amenity mismatch ($k < 1$), Government A is always able to undercut Government B by offering a lower tax to every firm $s \in [0, z]$.¹⁶ Government A maximizes rents by setting a limit tax. Because the cost of amenity mismatch is relatively low (for $k < 1$), the (lower) limit tax set by Government A is always enough to more than compensate for the larger degree of amenity mismatch.¹⁷

In the next result we show that if $k \geq 1$ then it is not possible for Government A to undercut Government B for all firms. Even if Government A sets taxes as low as possible, at $\tau_{As} = a$, a set of firms will still be better off locating in B . Therefore, when analyzing the case where $k \geq 1$, it will be helpful to re-introduce the notion of the marginal firm, \hat{s} , that is just indifferent between locating in either country. In the perfect tax discrimination game, the definition must be altered to allow for the fact that firms face individualized taxes:

$$\hat{s}(\tau_{As}, \tau_{Bs}) = \frac{\tau_{Bs} - \tau_{As}}{2k} + \frac{(z - b + a)}{2}.$$

The outcome in Stage 2 of the perfect tax discrimination game with costs of amenity mismatch relatively high are characterized in the following proposition.

Proposition 4. *Consider Stage 2 of a perfect tax discrimination game, with a and b fixed on the interval $[0, z]$. Assume A1 holds and that $a + b \leq z$, $a \geq 0$, $b \geq 0$. If $k \geq 1$ then for $\varepsilon > 0$ sufficiently small there exists a unique Nash equilibrium in taxes for this stage of the perfect tax discrimination game. A unique Nash equilibrium in taxes for each firm $s \in [0, z]$ is determined by the following taxes:*

¹⁶Note that this possibility of undercutting depends on the existence of a sufficiently small monetary unit. As a gets arbitrarily close to $z - b$, the smallest monetary unit must become arbitrarily small. But for given a and b , such a smallest monetary unit (ε) can always be found.

¹⁷The value of ε must be small enough so that Government A can set a tax τ_{As} sufficiently low and still make positive rent $\tau_{As} - a$. An explicit upper bound for the smallest money unit $\varepsilon \in (0, \hat{\varepsilon})$, where $\hat{\varepsilon} = (1 - k)(z - a - b)/2$, is established in the proof.

if $a + b = z$, then

$$\tau_{As}^* = \tau_{Bs}^* = a = z - b, \text{ for } a + b = z \text{ and } s \in [0, z];$$

if $a + b < z$, then

$$\begin{aligned} \tau_{As}^* &= a, \tau_{Bs}^* = z - b \text{ for } s = \hat{s}, \\ \tau_{As}^* &= \tau_{As}^{\lim}(\varepsilon), \tau_{Bs}^* = z - b, \text{ for } s \in [0, \hat{s}), \\ \tau_{As}^* &= a, \tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon) \text{ for } s \in (\hat{s}, z]. \end{aligned}$$

Proposition 4 works in exactly the same way as Proposition 3, except that Government B is able to limit tax the firms that are towards the upper end of $[0, z]$. Because the cost of mismatch is relatively high, firms towards the upper end of $[0, z]$ find it profitable to locate in B even when Government A sets its lowest possible tax $\tau_{As}^* = a$. Government B maximizes the rents that it extracts from them by setting a limit tax. In fact, Proposition 3 can be thought of as a special case of Proposition 4. In general, we should expect some firms to locate in each country. It is only when costs of amenity mismatch are below $k = 1$ that the government providing the amenity at a relatively low level can undercut the other government to such an extent that it attracts all firms.

Taking Propositions 3 and 4 together, we have seen that a Nash equilibrium exists for all possible values of k in Stage 2 of the perfect tax discrimination game. We close this subsection by making the observation formal.

Corollary 1. *Consider Stage 2 of a perfect tax discrimination game, with a and b fixed on the interval $[0, z]$. Assume A1 holds and that $a + b \leq z$, $a \geq 0$, $b \geq 0$ and $\varepsilon > 0$ sufficiently small. There exists a Nash equilibrium in taxes of the perfect tax discrimination game.*

2.4.2 Stage 1: Location

We now solve for Stage 1, defining an equilibrium in locations. Let T_A and T_B be tax schedules for Jurisdictions A and B respectively and let $r_A(T_A, T_B) = \int_{s \in [0, z]} r_{As}(\tau_{As}, \tau_{Bs})$ and $r_B(T_A, T_B) = \int_{s \in [0, z]} r_{Bs}(\tau_{As}, \tau_{Bs})$ be the corresponding overall rent functions. Using the equilibrium values $\tau_{As}^* = \tau_A^*(a, b, s, \varepsilon; k, z)$ and $\tau_{Bs}^* = \tau_B^*(a, b, s, \varepsilon; k, z)$ that we derived for Stage 2, the overall rent function for Government A can be written

$$r_A(T_A^*(a, b, s, \varepsilon; k, z), T_B^*(a, b, s, \varepsilon; k, z)) = r_A(a, b, s, \varepsilon; k, z).$$

Similarly, the overall rent function for Government B can be written

$$r_B(T_A^*(a, b, s, \varepsilon; k, z), T_B^*(a, b, s, \varepsilon; k, z)) = r_B(a, b, s, \varepsilon; k, z).$$

A location a^* of Government A is a *best reply* against a location b when it maximizes $r_A(a, b, s, \varepsilon; k, z)$ on the whole of \mathbb{R}_+ . A location b^* of Government B is a *best reply* against a location a when it maximizes $r_B(a, b, s, \varepsilon; k, z)$ on the whole of \mathbb{R}_+ . A *Nash equilibrium in locations* is a pair (a^*, b^*) such that a^* is a best reply to b^* and vice-versa.

First we characterize equilibrium when the cost of amenity mismatch is relatively low; that is, $k < 1$.

Proposition 5. *If $k < 1$ and $\varepsilon > 0$ sufficiently small then there exists a unique subgame perfect Nash equilibrium in pure strategies of the perfect tax discrimination game. Equilibrium is characterized by the point $a^* = 0$, $b^* = z$.*

In the unique equilibrium, neither government provides any amenities.¹⁸ To see the significance of this result, first recall that in the more familiar setting of perfect price discrimination by (private goods producing) firms, costs are exogenously given and in equilibrium, the price of the last unit sold is equal to its marginal

¹⁸Recall that b measures the distance from z , so when $b^* = z$ and $a^* = 0$ then both governments provide no amenities.

cost (limit pricing) and so the outcome is efficient. A firm's profit is equivalent to its contribution to social welfare, so profit maximization is equivalent to social welfare maximization. But in our model, governments' costs are endogenously determined by their location. From any position where governments are providing a positive level of amenities, Government A makes positive rents by attracting all firms to its jurisdiction while Government B makes zero rents (Proposition 3). Therefore, no government wants to be in the position of Government B . Each government has a unilateral incentive to undercut the other by reducing the level of amenity provision, in turn reducing taxes and attracting all firms to its jurisdiction. Because costs of amenity mismatch are relatively low, any firm can be more than compensated for amenity mismatch through lower taxation. Hence we have a 'race to the bottom' in tax rates and public good provision.

We now move on to consider the situation where amenity mismatch has a 'large' impact on costs; that is, $k \geq 1$. From Proposition 4 we saw that if $k \geq 1$ then, given a and b , some firms locate in each jurisdiction in the equilibrium of Stage 2. We now use the equilibrium taxes from Proposition 4 to solve overall rent functions in locations a and b for Stage 1. The overall rent function $r_A(a, b, s, \varepsilon; k, z)$ is shown to be strictly concave in a and the overall rent function $r_B(a, b, s, \varepsilon; k, z)$ is shown to be strictly concave in b . So from these we obtain candidates for equilibrium points a^* and b^* of Stage 1 of the game in the usual way. But these candidate points are based on the assumption that $a < z - b$. As we shall see, Proposition 6 shows that although b^* maximizes overall rents given $a < z - b$, Government B can make higher rents by setting $z - b \leq a$, presenting an incentive to deviate and undermining existence of equilibrium.

Assume $z - b > a$. Let $a^* \in \arg \max_a r_A(a, b, s, \varepsilon; k, z)$ and $b^* \in \arg \max_b r_B(a, b, s, \varepsilon; k, z)$. Using τ_{As}^* and τ_{Bs}^* from Proposition 4, note that

$$r_A(a, b, s, \varepsilon; k, z) = \int_{s \in [0, z]} r_{As}(\tau_{As}^*, \tau_{Bs}^*) = (a + (\hat{s} - a)/2)(1 + k)(z - a - b).$$

Taking the first derivative and solving for a yields a candidate for a^* :

$$a(b, k, z) = \frac{(k-1)(z-b)}{3k-1}.$$

Observe that for $k \geq 1$ the second derivative is negative $-\partial r_A / \partial a^2 = \frac{1}{2} \left(\frac{1}{k} - 2 - 3k \right) < 0$. So the objective function is concave. Again, from Proposition 4,

$$r_B(a, b, s, \varepsilon; k, z) = \int_{s \in [0, z]} r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = (b + (z - b - \hat{s})/2)(k-1)(z-a-b).$$

Taking the first derivative and solving for b yields a candidate for b^* :

$$b(a, k, z) = \frac{(1+k)(z-a)}{3k+1}.$$

Taking the second derivative, $\partial r_B / \partial b^2 = \frac{1}{2} \left(2 + \frac{1}{k} - 3k \right) \leq 0$ for $k \geq 1$. So the objective function is concave (weakly for $k = 1$). The functions $a(b, k, z) = (k-1)(z-b)/(3k-1)$ and $b(a, k, z) = (1+k)(z-a)/(3k+1)$ are reaction functions and can be solved for simultaneously to obtain a unique crossing point:

$$\begin{aligned} a(k, z) &= \frac{(k-1)z}{4k} \text{ and} \\ b(k, z) &= \frac{(k+1)z}{4k}. \end{aligned}$$

At the points $a(k, z) = (k-1)z/4k$, $b(k, z) = (k+1)z/4k$, each government maximizes its rent, taking as given the location of the other. But also notice that in solving this problem it has been assumed that $a < z - b$. Indeed, $a(k, z) = \frac{(k-1)z}{4k} < \frac{(3k-1)z}{4k} = z - b(k, z)$. But to establish that this is indeed an equilibrium, it must be checked that Government B does not have an incentive to adopt a level of amenity provision $(z - b) \leq a$. It is through the recognition of the possibility that Government B may have an incentive to deviate by setting $(z - b) \leq a$ that we obtain the following surprising result:

Proposition 6. *If $k \geq 1$ then there exists no subgame perfect Nash equilibrium in pure strategies of the perfect tax discrimination game.*

The intuition behind the result is as follows. At $a(k, z) = \frac{(k-1)z}{4k}$, $z - b(k, z) = \frac{(k+1)z}{4k}$, Government A makes higher rents than Government B . The difference in

rents when the Governments locate at these positions, and then adopt best response taxes in the second stage is $\frac{z^2}{4}$ in Government A 's favour. But because A does so much better, Government B has an incentive to deviate from $b(k, z) = \frac{(k+1)z}{4k}$ by locating in the same position as Government A , $a(k, z) = \frac{(k-1)z}{4k}$, and setting taxes slightly lower than Government A . (Thus B gives some of the additional surplus $\frac{z^2}{4}$ back to firms in exchange for relocation to B .) Jurisdiction B does not need to worry about losing the firms that, prior to the deviation, located in B because Government B makes more rents from the firms lured away from A . And prior to the deviation, B made zero rents from the firms that it now lures away from A . Thus, the rents that Government B makes under such a deviation are a net gain. This deviation contradicts equilibrium. Moreover, an equilibrium fails to exist because, from any position where $a \neq a(k, z)$, $b \neq b(k, z)$, there would be an incentive to move to these positions. And from these positions there is still an incentive to deviate, as just described. So no equilibrium can exist.

In the light of Corollary 1, the non-existence of equilibrium shown in Proposition 6 comes as a surprise. Corollary 1 shows that an equilibrium exists for all k . However, in Stage 2 of the game a and b are taken as fixed. In addition, it is assumed that $z - b \geq a$. The failure of equilibrium to exist comes about because a government positioned at $z - b$ on the interval has an incentive to deviate by setting a level of amenity provision equal to a and then undercut Government A on the tax. Then Government A has an incentive to deviate itself by changing its location. This possibility could not be accounted for in Stage 2 when locations were taken as fixed.

2.5 Efficiency

A standard social loss function is used to examine the efficiency implications of equilibrium (when it exists) under the respective regimes. The social loss function

is of the form

$$L = \int_{s \in [0, \hat{s}]} k |s - a| ds + \int_{s \in (\hat{s}, z]} k |z - b - s| ds.$$

This function aggregates the loss of potential profits that result from the divergence between amenity provision by each government and the ideal level of each firm.

Proposition 2 shows that a unique subgame perfect Nash equilibrium exists under the uniform tax game if and only if $0 < k \leq \frac{1}{7}$, and that the point $a^* = 0$, $b^* = 0$ is the equilibrium. Proposition 5 shows that a unique subgame perfect Nash equilibrium exists under the perfect tax discrimination game if $0 < k < 1$, and that the point $a^* = 0$, $b^* = 0$ is the equilibrium. To facilitate a comparison of efficiency across the two regimes, we assume that $0 < k \leq \frac{1}{7}$. Denote social loss under uniform taxation and perfect tax discrimination as L_u and L_p respectively. Then substituting equilibrium values and integrating it is immediate to see that

$$L_u = \left(\frac{1}{2}\right)^2 kz^2 < \frac{1}{2}kz^2 = L_p.$$

So under conditions where equilibrium would exist in both regimes, perfect tax discrimination brings about a lower level of social efficiency than uniform taxation under Hotelling amenity/tax competition. These solutions can be compared with the socially efficient outcome of $L^* = \frac{1}{8}kz^2$, which occurs when $a = b = \frac{z}{4}$.

2.6 Conclusions

This chapter seeks an explanation of why competition between governments fails to promote efficiency. The explanation we propose builds on Hotelling's observation that when firms compete not just over prices but over product characteristics, and when consumers' preferences over product characteristics vary, then efficiency is not promoted by competition. In the policy setting of the present chapter, competition between (Leviathan) governments fails to promote efficiency when governments compete over levels of amenity provision as well as taxes, and where firms' preferences for the level of amenity provision vary.

In the uniform tax game, when an equilibrium exists one government provides the amenity at a maximal level, which is inefficiently high, whilst the other government provides no amenity at all, which is inefficiently low. This result is driven by the variation in firms' ideal level of amenity provision. Then competition pushes governments 'too far' in opposite directions, rather than bringing about a universal race to the bottom or efficiency, the two outcomes on which most of the previous literature has focused.

The equilibrium that we demonstrate for uniform taxation appears to fit with recent empirical evidence, which shows persistent differences in levels of taxation and public good provision in areas where greater convergence had been expected. One example is in Europe, where a core and periphery has emerged despite significant efforts to avoid such an outcome. The core tends to be characterized by governments that tax and provide public amenities at a significantly higher level than in the periphery.

Interpreted more broadly, the equilibrium outcome may help to understand why aspects of economic development or legal reform may actually work against a government's (rent seeking) interests. A government in a country where public good provision is reckoned to be sub-optimally low may encounter resistance to reform. It has difficulties raising taxation because of resistance from both domestic and foreign firms whose original decision to locate or remain in that country was based on relatively low levels of amenity provision and taxation. An interesting thing about our analysis is that the usual presumption of downward pressure on developed country taxes and public good provision resulting from intergovernmental competition for firms does not follow. In this sense our theoretical predictions accord with the observation of a high-tax high-amenity providing core and low-tax low-amenity providing periphery of Europe. Our framework could similarly be used to help understand differences in amenity provision between the developed and developing worlds.

The failure of equilibrium to exist is taken as a metaphor for intense tax com-

petition. When the level of amenity provision offered by governments is similar then the weight of competition falls on tax levels. In the limit, because there is very little to choose between the two governments in terms of amenity levels, each government can attract all firms to its jurisdiction by undercutting the other with a small reduction in the tax level. When the degree of amenity mismatch has a sufficiently large impact on firms' costs, making them relatively responsive to changes in levels of amenity provision, then the system never settles down to (subgame perfect Nash) equilibrium. The governments both have an incentive to offer similar levels of amenities in an effort not to lose firms to the other. From the view point of each government, there is no tax level at which the other government does not have an incentive to attract all firms by setting a tax that is slightly lower.

One way to circumvent the incentive for governments to undercut each other is for each to offer tailor made tax-amenity packages to firms. There is a widespread perception that tax breaks are used in a similar vein. We model this policy environment as a 'perfect tax discrimination game'. We show that under perfect tax discrimination the equilibrium existence issue is partially resolved but that efficiency is worse than under uniform tax discrimination. The price paid by governments for greater stability through 'head to head' competition for each firm is that, once again when equilibrium exists, each government can attract the firm in question by lowering taxes, resulting in a 'race to the bottom'. In equilibrium, no amenities are provided by either government. As with uniform taxation, though, when the degree of amenity mismatch has a sufficiently large impact on firms' costs, making them relatively responsive to changes in levels of amenity provision, then the system never settles down to equilibrium. When no amenities are being offered, one government has an incentive to deviate by offering a level of amenity provision at a relatively high level. But when one government offers a positive level of amenities, then the other government can always do better by setting amenities at a slightly lower level and undercutting the first using taxes.

An alternative way to prevent intensive tax competition might lie with tax harmonization. For taxes set in the second stage governments could agree to set the same tax. Then the only issue would be in setting the level of amenity provision in the first stage. In the model of this present chapter it is clear that, given locations, under collusion the governments would have an incentive to raise taxes to the point where they had extracted all rents from firms. If perfect tax discrimination were possible then it is clear all rents would be extracted and the outcome would be efficient. Whilst economists might see such efficiency as an advantage, it is not clear that citizen-entrepreneurs would be happy to see all their profits transferred to politicians in the form of rents. Under uniform taxation the outcome is less obvious. Because of their differing requirements for amenity provision, firms make different profits. At a level of taxation where some firms could make positive profits and so a higher tax could extract further rents, other firms cannot make positive profits. The outcome would be dependent upon assumptions made about whether all firms must be profitable in equilibrium. The issue of tax harmonization within this framework is left to future research.

The framework of the present chapter is similar to a Tiebout model in that all firms can ‘vote with their feet’ for the jurisdiction that makes them better off (see Oates and Schwab 1988, Wooders 1989). So the inefficiencies that arise in the present model may seem surprising given that such mobility promotes efficiency in a Tiebout setting. The difference in outcomes appears to lie in the fact that in our setting there are just two jurisdictions whilst in a Tiebout setting there are many, combined with the fact that governments in a Tiebout setting are not Leviathans.

One might conjecture that increasing the number of jurisdictions in the model of this present chapter should bring about efficiency. On the face of it this appears to be true. To see why, assume that there are three governments in a uniform tax game. Introduce a third jurisdiction to the uniform tax game and assume that governments locate as far from each other as possible in amenity space, as in the equilibrium that we demonstrate for two jurisdictions. Computing social loss as

in Section 5, we find that the same efficient level of social loss is obtained as if two governments had located at a quarter and three quarters of the way along the interval. But it is far from clear that the three jurisdiction outcome is in fact an equilibrium. The government in the middle attracts half of all firms whilst the other two share a half. Therefore the other two may well have an incentive to deviate from such a situation. Given the discontinuities in the reaction function, it is not clear whether existence of equilibrium can be established in the three jurisdiction game.¹⁹

It is worth considering the implications of the present analysis for the public choice literature on tax competition, of which Besley and Smart (2001) is an example. In that literature, citizens are able to use yardstick competition to evaluate the performance of policy makers who may or may not be self-interested. Yardstick competition is shown to be a relatively effective mechanism in an environment where preferences for public good provision are uniform. If the level of public good provision in the other jurisdiction is higher than at home then there is evidence of under-performance by domestic politicians. It remains to be investigated whether the same holds in an environment where preferences for public good provision varies. One possibility would be to allow citizens to choose between a benevolent dictator and a Leviathan in a framework like the one of the present chapter, where agents' preferences for public good provision vary. It might then be possible to see whether Leviathan policy makers were induced to provide more efficient levels of public good provision or driven out of the policy arena all together. This seems like a promising area for further research.

¹⁹A larger number of agents has been introduced to a Hotelling framework by Salop (1979) where firms that compete for consumers are located on a circle. Note that such an approach would not be appropriate in our model because points on the interval denote levels of amenity provision rather than points in geographical space or time. So it does not make sense to join the two ends of the interval in order to form a circle.

2.7 Appendix

Lemma 1. Assume governments play a uniform tax game. For given τ_B , the unique tax that maximizes $r_A(\tau_A, \tau_B)$ on the firm sharing interval is

$$\tau_A(\tau_B; a, b, k, z) = k \left(\frac{a + \tau_B}{2k} + \frac{(z + a - b)}{2} \right).$$

For given τ_A , the unique tax τ_B that maximizes $r_B(\tau_A, \tau_B)$ on the firm sharing interval is

$$\tau_B(\tau_A; a, b, k, z) = k \left(\frac{(z - b) + \tau_A}{2k} + \frac{(z - a + b)}{2} \right).$$

If $\tau_A(\tau_B; a, b, k, z)$ and $\tau_B(\tau_A; a, b, k, z)$ are set simultaneously, then they can be solved for simultaneously to obtain:

$$\tau_A(a, b, k, z) = \frac{1}{3} (2a + (z - b) + (a - b)k + 3kz);$$

$$\tau_B(a, b, k, z) = \frac{1}{3} (2(z - b) + a + (b - a)k + 3kz).$$

Proof. To maximize rents over the firm sharing interval, Government A solves the problem

$$\max_{\tau_A} r_A(\tau_A, \tau_B) = (\tau_A - a) \hat{s}(\tau_A, \tau_B).$$

Expanding the objecting function using $\hat{s}(\tau_A, \tau_B) = (\tau_B - \tau_A)/2k + (z - b + a)/2$, we obtain

$$(\tau_A - a) \hat{s} = \frac{1}{2} (z + a - b) (\tau_A - a) - \frac{1}{2k} (\tau_A - a) \tau_A + \frac{1}{2k} (\tau_A - a) \tau_B.$$

Setting the first order condition equal to zero and rearranging obtains $\tau_A(\tau_B; a, b, k, z)$.

The second order condition is

$$\frac{\partial (r_A(\tau_A, \tau_B))}{\partial \tau_A} = -1/k,$$

so $r_A(\tau_A, \tau_B)$ must be strictly concave and $\tau_A(\tau_B; a, b, k, z)$ is the unique maximizer on the firm sharing interval.

Government B solves the analogous problem

$$\max_{\tau_B} r_B(\tau_A, \tau_B) = (\tau_B - (z - b))(z - \hat{s}(\tau_A, \tau_B)).$$

Expanding the objecting function, we obtain

$$\begin{aligned} (\tau_B - (z - b))(z - \hat{s}) = \\ \frac{1}{2}(z - a + b)(\tau_B - (z - b)) - \frac{1}{2k}(\tau_B - (z - b)) + \frac{1}{2k}(\tau_B - (z - b))\tau_A. \end{aligned}$$

Setting the first order condition equal to zero and rearranging obtains $\tau_B(\tau_A; a, b, k, z)$.

The second order condition once again is

$$\frac{\partial(r_B(\tau_A, \tau_B))}{\partial\tau_B} = -1/k,$$

so $r_B(\tau_A, \tau_B)$ must also be strictly concave and $\tau_B(\tau_A; a, b, k, z)$ is the unique maximizer on the firm sharing interval. \square

Lemma 2. *Under a uniform tax game, the tax $\tau_A(\tau_B; a, b, k, z)$ that maximizes $r_A(\tau_A, \tau_B)$ on the firm sharing interval $|\tau_A - \tau_B| \leq k(z - a - b)$ is a best response to τ_B if and only if, for any τ_B and $\varepsilon > 0$,*

$$r_A(\tau_A(\tau_B; a, b, k, z), \tau_B) \geq z(\tau_B - k(z - a - b) - a - \varepsilon).$$

Similarly, the tax $\tau_B(\tau_A; a, b, k, z)$ that maximizes $r_B(\tau_A, \tau_B)$ on the firm sharing interval $|\tau_A - \tau_B| \leq k(z - a - b)$ is a best response to τ_A if and only if, for any τ_A and $\varepsilon > 0$,

$$r_B(\tau_A, \tau_B(\tau_A; a, b, k, z)) \geq z(\tau_A - k(z - a - b) - (z - b) - \varepsilon).$$

Proof. For Government A , it is only necessary to check whether the tax $\tau_A = \tau_B - k(z - a - b) - \varepsilon$ yields a higher rent than $\tau_A = \tau_A(\tau_B; a, b, k, z)$; the rent maximizing tax on the firm sharing interval. By Lemma 1, $\tau_A(\tau_B; a, b, k, z) > 0$ and by construction $\hat{s} > 0$, so $r_A(\tau_A(\tau_B; a, b, k, z), \tau_B) > 0$. Therefore, the

alternative of setting $\tau_A = \tau_B + k(z - a - b)$, which yields zero rents, cannot yield higher rents than setting $\tau_A = \tau_A(\tau_B; a, b, k, z)$.

A parallel argument holds for Government B .

Having ruled out $\tau_A = \tau_B + k(z - a - b)$ as a strategy for Government A , sufficiency is immediate by definition of a best response. The tax $\tau_A = \tau_A(\tau_B; a, b, k, z)$ yields a rent $r_A(\tau_A(\tau_B; a, b, k, z), \tau_B)$, while the tax $\tau_A = \tau_B - k(z - a - b) - \varepsilon$ yields a rent $z(\tau_B - k(z - a - b) - a - \varepsilon)$. If $r_A(\tau_A(\tau_B; a, b, k, z), \tau_B) \geq z(\tau_B - k(z - a - b) - a - \varepsilon)$ then by definition $\tau_A = \tau_A(\tau_B; a, b, k, z)$ is a best response. Conversely, if to the contrary, $r_A(\tau_A(\tau_B; a, b, k, z), \tau_B) < z(\tau_B - k(z - a - b) - a - \varepsilon) = r_A(\tau_A, \tau_B)$ for some tax $\tau_A = \tau_B - k(z - a - b) - \varepsilon$, then by definition $\tau_A(\tau_B; a, b, k, z)$ cannot be a best response to τ_B . A parallel argument holds for Government B . \square

Proposition 1. *Assume governments play a uniform tax game, and that a and b are fixed on the interval $[0, z]$, with $a + b \leq z$, $a \geq 0$, $b \geq 0$. For $a + b = z$, both governments are at the same location and there always exists an equilibrium in which $\tau_A^* = a$, $\tau_B^* = z - b$.*

For $a + b < z$ there exists an equilibrium point if and only if the two following conditions hold:

$$(C1): r_A(\tau_A^*(\tau_B^*; a, b, k, z), \tau_B^*) \geq z(\tau_B^* - k(z - a - b) - a - \varepsilon) \Leftrightarrow \frac{((a - b)k + (z - a - b) + 3kz)^2}{18k} \geq \frac{z(2(a + 2b)k + 2(z - a - b) - 3\varepsilon)}{3}$$

$$(C2): r_B(\tau_B^*(\tau_A^*; a, b, k, z), \tau_A^*) \geq z(\tau_A^* - k(z - a - b) - (z - b) - \varepsilon) \Leftrightarrow \frac{((b - a)k - (z - a - b) + 3kz)^2}{18k} \geq \frac{z(2(2a + b)k - 2(z - a - b) - 3\varepsilon)}{3}$$

Whenever it exists, an equilibrium point is determined uniquely by the taxes

$$\begin{aligned} \tau_A^*(a, b; k, z) &= \frac{1}{3}(2a + (z - b) + (a - b)k + 3kz); \\ \tau_B^*(a, b; k, z) &= \frac{1}{3}(2(z - b) + a + (b - a)k + 3kz). \end{aligned}$$

Proof. For $a + b = z$ both governments are located in the same place and we effectively have a standard Bertrand equilibrium in homogeneous products.

Consider the case where $a + b < z$. Following d'Aspremont et al (1979), begin by showing that any equilibrium must satisfy the condition $|\tau_A^* - \tau_B^*| < k(z - a - b)$. Suppose that on the contrary, $|\tau_A^* - \tau_B^*| > k(z - a - b)$. Then the government that charges the strictly higher tax gets zero rents and gains by charging a tax equal to that of the other, contradicting the fact that (τ_A^*, τ_B^*) is an equilibrium.

Suppose then that $|\tau_A^* - \tau_B^*| = k(z - a - b)$. Take, for example, the case where $\tau_A^* - \tau_B^* = k(z - a - b)$. If $\tau_B^* = 0$ then the rents of Government B are zero and it would make positive rents by charging $0 < \tau_B^* < \tau_A^* + k(z - a - b)$. If $\tau_B^* > 0$ then there are two cases to consider: (i) Either Government A gets all firms to locate in A , in which case Government B can obtain positive rents by reducing τ_B^* . So Government B has an incentive to deviate from τ_B^* ; a contradiction; Or Government A has only a share of all firms and is able to capture all of them and make larger rents by charging a slightly lower tax. Let $\bar{s} < z$ be given by $\hat{s} = (\tau_B - \tau_A + (z - b + a)k) / 2k$ for which $\tau_A^* = \tau_B^* + k(z - a - b)$, given τ_B^* . At τ_A^* , Government A makes rents $r_A(\tau_A^*, \tau_B^*) = \tau_A^* \bar{s}$. For $\tau_A = \tau_A^* - \varepsilon$, the government makes rents $r_A(\tau_A, \tau_B^*) = \tau_A z$. For $\tau_A = \tau_A^* - \varepsilon$, where $\varepsilon = \varepsilon(z - \bar{s})\tau_A^*/z > 0$ the government makes rents $\tau_A z = \tau_A^* \bar{s}$. So for all $0 < \varepsilon < \varepsilon(z - \bar{s})\tau_A^*/z$, it is the case that $r_A(\tau_A^* - \varepsilon, \tau_B^*) = \tau_A z > \tau_A^* \bar{s}$; a contradiction. The only remaining possibility is that equilibrium must satisfy $|\tau_A^* - \tau_B^*| < k(z - a - b)$.

By definition of the 'rent to office' functions $r_A(\tau_A, \tau_B)$ and $r_B(\tau_A, \tau_B)$, for any equilibrium (τ_A^*, τ_B^*) , τ_A^* must maximize $\frac{1}{2}(z + a - b)(\tau_A - a) - \frac{1}{2k}(\tau_A - a)\tau_A + \frac{1}{2k}(\tau_A - a)\tau_B$ in the firm sharing interval $(\tau_B - k(z - a - b), \tau_B + k(z - a - b))$. An equivalent condition must hold for τ_B^* .

By Lemma 1, the first order conditions of this problem yield

$$\begin{aligned}\tau_A^*(\tau_B; a, b, k, z) &= \frac{a + \tau_B}{2} + \frac{(z + a - b)k}{2} \\ \tau_B^*(\tau_A; a, b, k, z) &= \frac{(z - b) + \tau_A}{2} + \frac{(z - a + b)k}{2}\end{aligned}$$

As we have just proved that firm sharing is necessary for equilibrium, the simultaneous solutions $\tau_A^*(a, b, k, z)$ and $\tau_B^*(a, b, k, z)$ given in Lemma 1 provide the equilibrium taxes.

To establish conditions under which this pair (τ_A^*, τ_B^*) is indeed an equilibrium, it remains to check that τ_A^* maximizes $r_A(\tau_A, \tau_B)$ not just on the interval $(\tau_B - k(z - a - b), \tau_B + k(z - a - b))$ but on the whole of the domain \mathbb{R}_+ , and similarly for τ_B^* . For fixed a and b , if τ_A^* is to be an equilibrium strategy given τ_B^* , by Lemma 2 we must have that for any $\varepsilon > 0$,

$$\begin{aligned} r_A(\tau_A^*, \tau_B^*) &= (\tau_A^* - a) \hat{s} = \frac{((a - b)k + (z - a - b) + 3kz)^2}{18k} \\ &\geq z(\tau_B^* - k(z - a - b) - a - \varepsilon). \end{aligned}$$

Substituting for \hat{s} using $\hat{s}(\tau_A, \tau_B) = (\tau_B - \tau_A)/2k + (z - b + a)/2$ and simplifying, we obtain condition (C1). By symmetry, we get (C2).

To show that (C1) and (C2) are also sufficient for (τ_A^*, τ_B^*) to be an equilibrium it remains only to check that they imply $|\tau_A^* - \tau_B^*| \leq k(z - a - b)$. This completes the proof of our proposition. \square

Proposition 2. *There exists a unique subgame perfect Nash equilibrium in pure strategies of a uniform tax game if and only if $0 < k \leq \frac{1}{7}$. If such an equilibrium exists then it is characterized (uniquely) by the point $a^* = b^* = 0$.*

Proof. Write $r_A(a, b; k, z)$ as $r_A(a, b)$ and $r_B(a, b; k, z)$ as $r_B(a, b)$ because k and z are held constant throughout.

First assume $0 < k \leq \frac{1}{7}$.

Suppose that the pair (a^*, b^*) is a Nash equilibrium, where *either* a is interior or b is interior (or both); $a \in (0, z)$ or $b \in (0, z)$. Take b^* as given and let $a^* \in (0, z)$. But by Lemma 3, $\partial^2 r_A(a^*, b^*)/\partial a^2 = (k - 1)^2/9k > 0$. If $\partial r_A(a^*, b^*)/\partial a > (<) 0$ then rents can be increased by increasing (decreasing) a , contradicting equilibrium. If $\partial r_A(a^*, b^*)/\partial a = 0$ then rents can be increased either by increasing or by decreasing a , again contradicting equilibrium. The same argument can be made for $b^* \in (0, z)$, holding a^* constant, as $\partial^2 r_B(a, b)/\partial b^2 = (k + 1)^2/9k > 0$.

Therefore, the only candidates for an equilibrium pair are the corner solutions $(a^*, b^*) = (0, 0)$, $(0, z)$ and $(z, 0)$ (noting that (z, z) violates $a + b \leq z$). The three cases are taken in order. First we show why $(a^*, b^*) = (0, 0)$ is an equilibrium. First observe that $\partial r_A(a, b) / \partial a = (1 - k)((1 - k)a + (1 + k)b - (1 + 3k)z) / 9k$. Using $b^* = 0$, $\partial r_A(a, b^*) / \partial a = (1 - k)((1 - k)a - (1 + 3k)z) / 9k < 0$ for all $a \in [0, z]$. To see this, note that even when a takes its largest positive value at $a = z$, $\partial r_A(a, b^*) / \partial a = -4(1 - k) / 9z < 0$. Thus we have a corner solution. Rents could be increased were it possible to reduce a below the level $a = 0$. But this is not possible so $a^* = 0$ is a best response to $b^* = 0$.

Now take $a^* = 0$ as given and observe that $\partial r_B(a^*, b^*) / \partial b = (1 + k)((1 + k)b + z(3k - 1)) / 9k$. If $b = 0$ then $\partial r_B(a^*, b) / \partial b = (1 + k)(3k - 1) / 9k < 0$. But if $b = z$ then $\partial r_B(a^*, b) / \partial b = 4(1 + k)z / 9 > 0$. So both $b = 0$ and $b = z$ could in principle be stable corner solutions (see from above that the second order condition is satisfied). The matter of which is a best response depends upon which yields the higher rent; $r_B(0, 0) = ((3k - 1)z)^2 / 18k$ or $r_B(0, z) = 8kz^2 / 9$. Solving $r_B(0, 0) = r_B(0, z)$ in terms of k we find that $k = \frac{1}{7}$. It is then easy to see that $r_B(0, 0) \geq r_B(0, z)$ for $0 < k \leq \frac{1}{7}$, with $r_B(0, 0) > r_B(0, z)$ for $0 < k < \frac{1}{7}$. So $b^* = 0$ is a best response to $a^* = 0$. Therefore, $(a^*, b^*) = (0, 0)$ is a Nash equilibrium in locations.

Next suppose $(a^*, b^*) = (0, z)$ is a Nash equilibrium in locations. But then (C1) fails;

$$\frac{((a - b)k + (z - a - b) + 3kz)^2}{18k} - \frac{z(2(a + 2b)k + 2(z - a - b))}{3} = -\frac{10kz^2}{9};$$

Next suppose $(a^*, b^*) = (z, 0)$ is a Nash equilibrium in locations. But then (C2) fails;

$$\frac{((b - a)k - (z - a - b) + 3kz)^2}{18k} - \frac{z(2(2a + b)k - 2(z - a - b))}{3} = -\frac{10kz^2}{9};$$

a contradiction.

Now assume $k > \frac{1}{7}$.

Suppose that $(a^*, b^*) = (0, 0)$ is a Nash equilibrium in locations. But $r_B(0, z) >$

$r_B(0,0)$. So there is a unilateral incentive for Government B to deviate from $b^* = 0$; a contradiction.

The solutions $(a^*, b^*) = (0, z)$ and $(a^*, b^*) = (z, 0)$ can be ruled out as candidates for a Nash equilibrium in locations for the same reason as when $0 < k \leq \frac{1}{7}$; Conditions (C1) and (C2) fail in the respective cases. \square

Lemma 4. *Consider a perfect tax discrimination game and assume A1 holds. Fix a and b so that $z - b > a$.*

If, for some firm $s \in [0, z]$, $a < \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then for $\varepsilon > 0$ sufficiently small Government A 's unique best response is $\tau_{As}^ = \tau_{As}^{\lim}(\varepsilon)$. If $a \geq \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then $\tau_{As}^* = a$ is a best response for Government A .*

If, for some firm $s \in [0, z]$, $z - b < \tau_{As} + k(|s - a| - |(z - b) - s|)$ then for $\varepsilon > 0$ sufficiently small Government B 's unique best response is $\tau_{Bs}^(\varepsilon) = \tau_{Bs}^{\lim}$. If $z - b \geq \tau_{As} + k(|s - a| - |(z - b) - s|)$ then $\tau_{Bs}^* = z - b$ is a best response for Government B .*

Proof. It is assumed that $\varepsilon > 0$ and arbitrarily small. The exact bound on ε is established below.

Consider Government A 's best response first. Fix a, b and τ_{Bs} so that $a < \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ and suppose to the contrary that $\tau_{As}^{\lim} = \tau_{Bs} + k(|(z - b) - s| - |s - a|) - \varepsilon$ is not the unique best response. Then by definition, there must be some other tax that yields a higher rent. First suppose that the best response tax is lower than τ_{As}^{\lim} , obtained by setting $\varepsilon' > \varepsilon$. Write $r_{As}(\varepsilon)$ for the rent obtained from setting tax $\tau_{As} = \tau_{Bs} + k(|(z - b) - s| - |s - a|) - \varepsilon$. Taking the difference in rents we obtain $r_{As}(\varepsilon') - r_{As}(\varepsilon) = -\varepsilon' + \varepsilon < 0$. So rents are lower under a lower tax; contradiction.

Next suppose that the best response tax is higher than τ_{As}^{\lim} . Suppose that Government A raises the tax by the smallest possible amount, to $\tau_{As} = \tau_{Bs} + k(|(z - b) - s| - |s - a|)$. Write the rent associated with this tax rate as $r_{As}(0)$. At this tax, the firm s is indifferent between the two jurisdictions. By A1, the firm

s locates in A with probability $\frac{1}{2}$. Taking the difference in rents we obtain

$$\begin{aligned} r_{As}(0) - r_{As}(\varepsilon) &= \frac{1}{2} (\tau_{Bs} + k(|(z-b) - s| - |s-a|) - a) \\ &\quad - (\tau_{Bs} + k(|(z-b) - s| - |s-a|) - \varepsilon - a) \\ &= -\frac{1}{2} (\tau_{Bs} + k(|(z-b) - s| - |s-a|) - a) + \varepsilon. \end{aligned}$$

But it is always possible to pick ε sufficiently small to ensure that $r_{As}(0) - r_{As}(\varepsilon) < 0$; contradiction.

By definition of $r_{As}(\tau_{As}, \tau_{Bs})$, if Government A sets a tax $\tau_{As} > \tau_{Bs} + k(|(z-b) - s| - |s-a|)$ then $r_{As}(\tau_{As}, \tau_{Bs}) = 0$, whilst $r_{As}(\tau_{As}^{\lim}(\varepsilon), \tau_{Bs}) > 0$. So rents are lower under a higher tax; contradiction. So we have established that if $a < \tau_{Bs} + k(|(z-b) - s| - |s-a|)$ then the unique best response is $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$.

Now fix a, b and τ_{Bs} so that $a \geq \tau_{Bs} + k(|(z-b) - s| - |s-a|)$ and suppose to the contrary that $\tau_{As}^* = a$ is not a best response. Then by definition there must be some other tax that yields a higher rent. First note that $r_{As}(\tau_{As}^*, \tau_{Bs}) = 0$. Clearly, $\tau_{As} < a$ would yield $r_{As}(\tau_{As}, \tau_{Bs}) < 0$; contradiction. Now suppose $\tau_{As} > a$. But then $\tau_{As} > \tau_{Bs} + k(|(z-b) - s| - |s-a|)$ and so, by definition of the rent function, $r_{As}(\tau_{As}, \tau_{Bs}) = 0$. So rents are not higher under a higher tax; contradiction.

An analogous set of arguments can be used to establish the corresponding results for the best response of Government B . \square

Proposition 3. *Consider Stage 2 of a perfect tax discrimination game, with a and b fixed on the interval $[0, z]$. Assume A1 holds and that $a + b \leq z$, $a \geq 0$, $b \geq 0$. If $k < 1$ then for ε sufficiently small there exists a unique Nash equilibrium in taxes for this stage of the perfect tax discrimination game. A unique Nash equilibrium in taxes for each firm $s \in [0, z]$ is determined by the following taxes:*

if $a + b = z$,

$$\tau_{As}^* = \tau_{Bs}^* = a = z - b;$$

if $a + b < z$,

$$\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon), \quad \tau_{Bs}^* = z - b.$$

Proof.

For $a + b = z$, both governments are located in the same place and we effectively have a standard Bertrand equilibrium in homogeneous products.

Consider the case where $a + b < z$. It is assumed that $\varepsilon > 0$ and arbitrarily small. An explicit upper bound $\bar{\varepsilon} = (1 - k)(z - a - b)/2$ for ε will be established in the proof below.

We will show that for all $s \in [0, z]$ the following pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium.

$$\begin{aligned} \tau_{As}^* &= \tau_{As}^{\lim}(\varepsilon) = z - b + k(|z - b - s| - |s - a|) - \varepsilon; \\ \tau_{Bs}^* &= z - b. \end{aligned}$$

First check the firm's location decision. We take the difference between the cost to locating in B and locating in A :

$$\begin{aligned} c_{Bs}(\tau_{Bs}^*) - c_{As}(\tau_{As}^*) &= \tau_{Bs}^* + k|z - b - s| - \tau_{As}^* - k|s - a| \\ &= z - b + k|z - b - s| - (z - b) \\ &\quad - (k(|z - b - s| - |s - a|) - \varepsilon) - k|s - a| \\ &= \varepsilon \end{aligned}$$

For each firm $s \in [0, z]$, profits made in Jurisdiction A are higher by ε than profits made in Jurisdiction B . Therefore, each firm locates in A .

To check that the pair $(\tau_{As}^*, \tau_{Bs}^*)$ does indeed represent a Nash equilibrium, suppose not. Then either $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is not a best response to $\tau_{Bs}^* = z - b$ or vice versa. First suppose that $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is not a best response to $\tau_{Bs}^* = z - b$. If $\tau_{Bs}^* = z - b$ then for all $k < 1$,

$$a < \tau_{Bs}^* + k(|(z - b) - s| - |s - a|).$$

To see this, note that for $s \geq z - b$, it is the case that $|(z - b) - s| - |s - a| = -|z - a - b|$ and for $s < z - b$ it is the case that $|(z - b) - s| - |s - a| > -|z - a - b|$. Using this and $\tau_{Bs}^* = z - b$ we have $0 < z - a - b + k(|(z - b) - s| - |s - a|)$. But by Lemma 4, if $a < \tau_{Bs}^* + k(|(z - b) - s| - |s - a|)$ then $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is a best response to $\tau_{Bs}^* = z - b$; a contradiction.

We now establish the upper bound $\bar{\varepsilon} = (1 - k)(z - a - b)/2$ on ε . Recall that Lemma 4 required ε to be sufficiently small as to ensure that $r_{As}(0) - r_{As}(\varepsilon) < 0$. Let $\bar{\varepsilon} = (1 - k)(z - a - b)/2$. If $\varepsilon < \bar{\varepsilon}$ then $r_{As}(0) - r_{As}(\varepsilon) < 0$ for all $s \in [0, z]$. To see why, use the fact that $(|(z - b) - s| - |s - a|) \geq -|z - a - b|$ and $\tau_{Bs}^* = z - b$ in the expression for $r_{As}(0) - r_{As}(\bar{\varepsilon})$:

$$\begin{aligned} r_{As}(0) - r_{As}(\bar{\varepsilon}) &= -\frac{1}{2}(z - a - b + k(|(z - b) - s| - |s - a|)) \\ &\quad + \frac{1}{2}(1 - k)(z - a - b) \\ &\leq -\frac{1}{2}(z - a - b - k|z - a - b|) \\ &\quad + \frac{1}{2}(1 - k)(z - a - b) \\ &= 0 \text{ for all } s \in [0, z]. \end{aligned}$$

It follows directly that if $\varepsilon < \bar{\varepsilon} = (1 - k)(z - a - b)/2$ then $r_{As}(0) - r_{As}(\bar{\varepsilon}) < 0$ for all $s \in [0, z]$. So there exists an $\bar{\varepsilon}$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$, $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is a best response to $\tau_{Bs}^* = z - b$.

Now suppose that $\tau_{Bs} = z - b$ is not a best response to $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$. If $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$ then

$$\begin{aligned} &\tau_{As}^* + k(|s - a| - |(z - b) - s|) \\ &= z - b + k(|z - b - s| - |s - a|) - \varepsilon \\ &\quad + k(|s - a| - |(z - b) - s|) \\ &= z - b - \varepsilon. \end{aligned}$$

But by Lemma 4, if $z - b \geq \tau_{As}^* + k(|s - a| - |(z - b) - s|)$ then $\tau_{Bs}^* = z - b$ is a best response to $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$; a contradiction. Note that this does not depend on the value of k and s .

We now demonstrate uniqueness of this Nash equilibrium. We already know from Lemma 4 that $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$ is the unique best response to $\tau_{Bs}^* = z - b$. On the other hand, $\tau_{Bs}^* = z - b$ is not a unique best response to $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$, as any $\tau_{Bs}^* \geq z - b$ earns $r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = 0$ for Government B . However, $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$, $\tau_{Bs} > z - b$ is not a Nash equilibrium. To see this, set some $\tau_{Bs} > z - b$ and $\tau_{As} = \tau_{As}^{\lim}(\varepsilon) = \tau_{Bs} + k(|z - b - s| - |s - a|) - \varepsilon$. As A is limit pricing the firm s , the firm locates in A and Government B makes rent $r_{Bs}(\tau_{As}, \tau_{Bs}) = 0$. As long as $\tau_{Bs}^{\lim}(\varepsilon) > z - b$, Government B has an incentive to deviate from τ_{Bs} by setting $\tau_{Bs}^{\lim}(\varepsilon)$, attracting the firm s to Jurisdiction B and making $r_{Bs}(\tau_{As}, \tau_{Bs}) > 0$. Only at $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$, $\tau_{Bs}^* = z - b$ does Government B not have a deviation that could make positive rents. In order to attract the firm s to B the government must set $\tau_{Bs} < z - b$ and this would violate condition (ii) of equilibrium.

As we have characterized a unique Nash equilibrium for all $s \in [0, z]$, we have demonstrated that there exists a unique Nash equilibrium in taxes. \square

Proposition 4. *Consider Stage 2 of a perfect tax discrimination game, with a and b fixed on the interval $[0, z]$. Assume A1 holds and that $a + b \leq z$, $a \geq 0$, $b \geq 0$. If $k \geq 1$ then for ε sufficiently small there exists a unique Nash equilibrium in taxes for this stage of the perfect tax discrimination game. A unique Nash equilibrium in taxes for each firm $s \in [0, z]$ is determined by the following taxes:*

if $a + b = z$, then

$$\tau_{As}^* = \tau_{Bs}^* = a = z - b, \text{ for } a + b = z \text{ and } s \in [0, z];$$

if $a + b < z$, then

$$\begin{aligned} \tau_{As}^* &= a, \tau_{Bs}^* = z - b \text{ for } s = \hat{s}, \\ \tau_{As}^* &= \tau_{As}^{\lim}(\varepsilon), \tau_{Bs}^* = z - b, \text{ for } s \in [0, \hat{s}), \\ \tau_{As}^* &= a, \tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon) \text{ for } s \in (\hat{s}, z]. \end{aligned}$$

Proof. For $a + b = z$, both governments are located in the same place and we effectively have a standard Bertrand equilibrium in homogeneous products.

Consider the case where $a + b < z$. It is assumed that $\varepsilon > 0$ and arbitrarily small.

First take the firm $s = \hat{s}$. To solve for its location, use $\tau_{As}^* = a$ and $\tau_{Bs}^* = z - b$ in $\hat{s} = (\tau_{Bs} - \tau_{As}) / 2k + (z - b + a) / 2$ to obtain

$$\hat{s} = \frac{(1+k)(z-b) + (k-1)a}{2k}.$$

It is straightforward to verify that $a < \hat{s} < z - b$ for $k > 1$, and that $\hat{s} \rightarrow z - b$ from below as $k \rightarrow 1$ (from above). By construction, $s = \hat{s}$ makes the same profits in Jurisdiction A as in Jurisdiction B . Therefore, by A1, the probability that it locates in each jurisdiction is $\frac{1}{2}$.

We will now show that the following pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium for $s = \hat{s}$;

$$\begin{aligned}\tau_{As}^* &= a; \\ \tau_{Bs}^* &= z - b.\end{aligned}$$

To check that the pair $(\tau_{As}^*, \tau_{Bs}^*)$ does indeed represent a Nash equilibrium for $s = \hat{s}$, suppose not. Then either $\tau_{As} = a$ is not a best response to $\tau_{Bs}^* = z - b$ or vice versa.

First suppose that $\tau_{As} = a$ is not a best response to $\tau_{Bs}^* = z - b$. For $\tau_{As} = a$, rents are given by

$$\begin{aligned}r_{As}(\tau_{As}, \tau_{Bs}^*) &= \tau_{As} - a \\ &= 0.\end{aligned}$$

Setting $\tau_{As} < a$ contradicts condition (ii) of equilibrium. If Government A deviates by setting $\tau_{As} > a$ then the firm makes higher profits by locating in Jurisdiction B , as a result of which $r_{As}(\tau_{As}, \tau_{Bs}^*) = 0$. So there exists no profitable deviation from $\tau_{As} = a$; contradiction. An analogous argument holds for $\tau_{Bs}^* = z - b$.

Next take firms in the interval $s \in [0, \hat{s})$. We will show that for all such firms

the following pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium.

$$\begin{aligned}\tau_{As}^* &= \tau_{As}^{\lim}(\varepsilon) = z - b + k(|z - b - s| - |s - a|) - \varepsilon; \\ \tau_{Bs}^* &= z - b;\end{aligned}$$

where it is assumed that $\varepsilon > 0$ and arbitrarily small.

First check the firm's location decision. We take the difference between the cost to locating in B and locating in A :

$$\begin{aligned}c_{Bs}(\tau_{Bs}^*) - c_{As}(\tau_{As}^*) &= \tau_{Bs}^* + k|z - b - s| - \tau_{As}^* - k|s - a| \\ &= z - b + k|z - b - s| \\ &\quad - (z - b) - (k(|z - b - s| - |s - a|) - \varepsilon) - k|s - a| \\ &= \varepsilon\end{aligned}$$

For each firm $s \in [0, \hat{s})$, profits made in Jurisdiction A are higher by ε than profits made in Jurisdiction B . Therefore, each firm locates in A .

To check that the pair $(\tau_{As}^*, \tau_{Bs}^*)$ does indeed represent a Nash equilibrium, suppose not. Then either $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is not a best response to $\tau_{Bs}^* = z - b$ or vice versa. First suppose that $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is not a best response to $\tau_{Bs}^* = z - b$. Check that Government A makes non-negative rents at $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$. Otherwise condition (ii) of equilibrium is violated. Note that we can represent any firm $s \in [0, \hat{s})$ as $s = \hat{s} - \delta > 0$, where $0 < \delta \leq \hat{s}$. Using this notation, we find that $r_{As} = 2k\delta - \varepsilon$. To see why, use $\hat{s} = ((1 + k)(z - b) + (k - 1)a)/2k$, $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ and $\tau_{Bs}^* = z - b$ in $r_{As} = \tau_{As} - a$. As $k, \delta > 0$, it is always possible to pick an ε sufficiently small to ensure that $r_{As} = 2k\delta - \varepsilon > 0$.

If $\tau_{Bs}^* = z - b$ then for $s \in [0, \hat{s})$,

$$a < \tau_{Bs}^* + k(|(z - b) - s| - |s - a|).$$

To see this, now use $s = ((1 + k)(z - b) + (k - 1)a)/2k - \delta$ and $\tau_{Bs}^* = z - b$ in the above expression to show that

$$\tau_{Bs}^* - a + k(|(z - b) - s| - |s - a|) = 2k\delta > 0.$$

But by Lemma 4, if $a < \tau_{Bs}^* + k(|(z - b) - s| - |s - a|)$ then $\tau_{As} = \tau_{As}^{\lim}(\varepsilon)$ is a best response to $\tau_{Bs}^* = z - b$; a contradiction.

Suppose that $\tau_{Bs} = z - b$ is not a best response to $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$. Exactly the same argument as in Proposition 3 is used to establish a contradiction. (Recall that the argument used in Proposition 3 was independent of the value of k and s). We have that, for all $s \in [0, \hat{s})$, the pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium.

We now demonstrate uniqueness of this Nash equilibrium. Once again, exactly the same argument as in Proposition 3 is used to establish that $\tau_{As}^* = \tau_{As}^{\lim}(\varepsilon)$, $\tau_{Bs}^* = z - b$ is unique.

Now consider all $s \in (\hat{s}, z]$. For such firms we will show that the following pair $(\tau_{As}^*, \tau_{Bs}^*)$ is a Nash equilibrium:

$$\begin{aligned}\tau_{As}^* &= a; \\ \tau_{Bs}^* &= \tau_{Bs}^{\lim}(\varepsilon) = a - k(|z - b - s| - |s - a|) - \varepsilon.\end{aligned}$$

First check the firm's location decision. We take the difference between the cost to locating in B and locating in A :

$$\begin{aligned}c_{Bs}(\tau_{Bs}^*) - c_{As}(\tau_{As}^*) &= \tau_{Bs}^* + k|z - b - s| - \tau_{As}^* - k|s - a| \\ &= a - k(|z - b - s| - |s - a|) - \varepsilon + k|z - b - s| - a - k|s - a| \\ &= -\varepsilon\end{aligned}$$

So costs are lower and therefore profits are higher for the firm if it locates in Country B .

To check that the pair $(\tau_{As}^*, \tau_{Bs}^*)$ does indeed represent a Nash equilibrium for $s \in (\hat{s}, z]$, suppose not. Then either $\tau_{As} = a$ is not a best response to $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$ or vice versa.

Suppose that $\tau_{As} = a$ is not a best response to $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$. If $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$

then

$$\begin{aligned}
 & \tau_{Bs}^* + k(|(z - b) - s| - |s - a|) \\
 = & a - k(|z - b - s| - |s - a|) - \varepsilon \\
 & + k(|(z - b) - s| - |s - a|) \\
 = & a - \varepsilon
 \end{aligned}$$

But by Lemma 4, if $a \geq \tau_{Bs} + k(|(z - b) - s| - |s - a|)$ then $\tau_{As}^* = a$ is a best response to $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$ for Government A; contradiction.

Now suppose that $\tau_{Bs} = \tau_{Bs}^{\lim}(\varepsilon)$ is not a best response to $\tau_{As}^* = a$. Check that Government B makes non-negative rents at $\tau_{Bs} = \tau_{Bs}^{\lim}(\varepsilon)$. Using the fact that $s = \hat{s} + \delta > 0$, where $\delta > 0$, we find that $r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = 2k\delta - \varepsilon$. To see this, use $\hat{s} = ((1 + k)(z - b) + (k - 1)a)/2k$ and $\tau_{Bs} = \tau_{Bs}^{\lim}(\varepsilon)$ in $r_{Bs} = \tau_{Bs} - z - b$. As $k, \delta > 0$, it is always possible to pick an ε sufficiently small to ensure that $r_{Bs} = 2k\delta - \varepsilon > 0$.

If $\tau_{As}^* = a$ then for $s \in (\hat{s}, z]$,

$$z - b < \tau_{As}^* + k(|s - a| - |(z - b) - s|)$$

To see this, use $s = ((1 + k)(z - b) + (k - 1)a)/2k - \delta$ and $\tau_{As}^* = a$ in the above expression to show that

$$\tau_{As}^* - (z - b) + k(|s - a| - |(z - b) - s|) = 2k\delta > 0.$$

But by Lemma 4, if $z - b < \tau_{As}^* + k(|s - a| - |(z - b) - s|)$ then $\tau_{Bs}^* = z - b$ is a best response to $\tau_{As}^* = a$ for Government B; contradiction.

We now demonstrate uniqueness of this Nash equilibrium. We already know from Lemma 4 that $\tau_{Bs}^* = \tau_{Bs}^{\lim}$ is the unique best response to $\tau_{As}^* = a$. On the other hand, $\tau_{As}^* = a$ is not a unique best response to $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$, as any $\tau_{As}^* \geq a$ earns $r_{As}(\tau_{As}^*, \tau_{Bs}^*) = 0$ for Government A. However, $\tau_{As} > a$, $\tau_{Bs} = \tau_{Bs}^{\lim}(\varepsilon)$ is not a Nash equilibrium. To see this, set some $\tau_{As} > a$ and $\tau_{Bs} = \tau_{Bs}^{\lim}(\varepsilon) = \tau_{As} - k(|z - b - s| - |s - a|) - \varepsilon$. As B is limit pricing the firm

s , the firm locates in B and Government A makes rent $r_{As}(\tau_{As}, \tau_{Bs}) = 0$. As long as $\tau_{As}^{\lim}(\varepsilon) > a$, Government A has an incentive to deviate from τ_{As} by setting $\tau_{As}^{\lim}(\varepsilon)$, attracting the firm s to Jurisdiction A and making $r_{As}(\tau_{As}, \tau_{Bs}) > 0$. Only at $\tau_{As}^* = a$, $\tau_{Bs}^* = \tau_{Bs}^{\lim}(\varepsilon)$ does Government A not have a deviation that could make positive rents. In order to attract the firm s to A the government must set $\tau_{As} < a$ and this would violate condition (ii) of equilibrium.

As we have characterized a unique Nash equilibrium for all $s \in [0, z]$, we have demonstrated that there exists a unique Nash equilibrium in taxes. \square

Proposition 5. *If $k < 1$ and $\varepsilon > 0$ sufficiently small then there exists a unique subgame perfect Nash equilibrium in pure strategies of the perfect tax discrimination game. Equilibrium is characterized by the point $a^* = 0$, $b^* = z$.*

Proof. We assume that $\varepsilon > 0$ and arbitrarily small. First we show that, for $a^* = 0$ and $b^* = z$, a^* is a best response to b^* and vice-versa.

Look for B 's incentive to deviate. At $b^* = z$, ($a^* = 0$), $|(z - b) - s| - |s - a| = 0$ for all $s \in [0, z]$. So $\tau_{Bs}^* = z - b = 0$ and $r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = \tau_{Bs}^* - (z - b) = 0$ for all s . Government B cannot deviate by raising b , so the only option would be to deviate by lowering b . But by Proposition 3, it follows that if Government B sets $b < z$ so that $\tau_{Bs} = z - b > 0$, whilst Government A sets $\tau_{As}^* = a = 0$, then all firms make higher profits by locating in Jurisdiction A ; and so $r_{Bs}(\tau_{As}^*, \tau_{Bs}) = 0$ for all s . So there is no profitable deviation for B . By symmetry, A has no incentive to deviate by raising a .

Now suppose that some other equilibrium exists where $a \in (0, z]$ and $b \in [0, z)$ and $z - b \geq a$. If $z - b = a$ then $\tau_{As}^* = \tau_{Bs}^* = z - b$ and $r_{As}(\tau_{As}^*, \tau_{Bs}^*) = r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = 0$ for all s . But by Proposition 3, Government A could attract all firms by lowering a and make positive rents. By symmetry, Government B has an incentive to raise b to a point where $z - b < a$ and attract all firms in order to make positive rents. Therefore, no values $a \in (0, z]$ and $b \in [0, z)$ can be an equilibrium. \square

Proposition 6. *If $k \geq 1$ then there exists no subgame perfect Nash equilibrium*

in pure strategies of the perfect tax discrimination game.

Proof. First let $z - b > a$. Let $a^* \in \arg \max_a r_A(a, b, s, \varepsilon; k, z)$ and $b^* \in \arg \max_b r_B(a, b, s, \varepsilon; k, z)$. Using τ_{As}^* and τ_{Bs}^* from Proposition 4,

$$r_A(a, b, s, \varepsilon; k, z) = \int_{s \in [0, z]} r_{As}(\tau_{As}^*, \tau_{Bs}^*) = (a + (\hat{s} - a)/2)(1 + k)(z - a - b).$$

Taking the first derivative and solving for a yields a candidate for a^* :

$$a(b, k, z) = \frac{(k - 1)(z - b)}{3k - 1}.$$

Recall that, for $k \geq 1$, $\partial r_A / \partial a^2 = \frac{1}{2} \left(\frac{1}{k} - 2 - 3k \right) < 0$. So the objective function is concave.

Again, from Proposition 4,

$$r_B(a, b, s, \varepsilon; k, z) = \int_{s \in [0, z]} r_{Bs}(\tau_{As}^*, \tau_{Bs}^*) = (b + (z - b - \hat{s})/2)(k - 1)(z - a - b).$$

Taking the first derivative and solving for b yields a candidate for b^* :

$$b(a, k, z) = \frac{(1 + k)(z - a)}{3k + 1}.$$

Recall once again that $\partial r_B / \partial b^2 = \frac{1}{2} \left(2 + \frac{1}{k} - 3k \right) \leq 0$ for $k \geq 1$. So the objective function is concave (weakly for $k = 1$). Solving $a(b, k, z)$ and $b(a, k, z)$ simultaneously for a and b in terms of parameters k and z we have

$$\begin{aligned} a(k, z) &= \frac{(k - 1)z}{4k} \text{ and} \\ b(k, z) &= \frac{(k + 1)z}{4k}. \end{aligned}$$

At the points $a(k, z) = (k - 1)z/4k$, $b(k, z) = (k + 1)z/4k$, each government maximizes its rent.

Now suppose, contrary to the statement of the proposition, that there exists a subgame perfect equilibrium of this game. Then given the global concavity of the payoff functions $r_A(a, b, s, \varepsilon; k, z)$ and $r_B(a, b, s, \varepsilon; k, z)$, equilibrium must be characterized by the points following points:

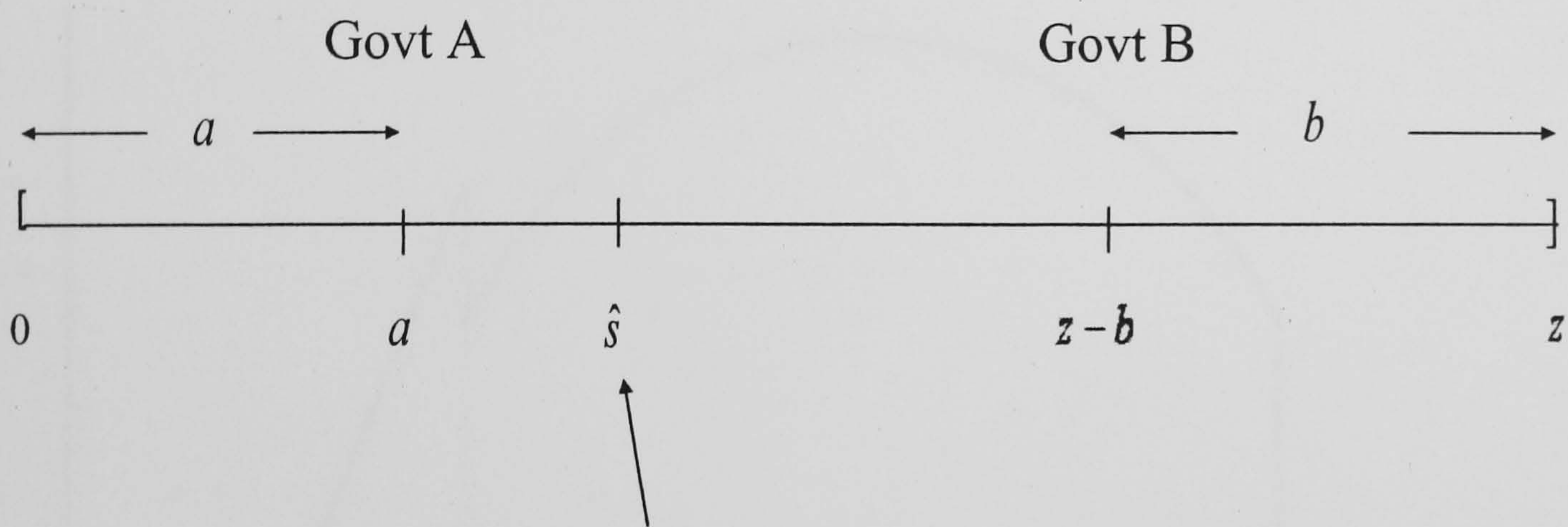
$$\begin{aligned} a^* &= \frac{(k - 1)z}{4k}; \\ b^* &= \frac{(k + 1)z}{4k}. \end{aligned}$$

Now using these values calculate the difference in rents to governments a and b :

$$r_A(a^*, b^*, s, \varepsilon; k, z) - r_B(a^*, b^*, s, \varepsilon; k, z) = \frac{z^2}{4}.$$

Therefore, for any location $z - b > a$ chosen by Government B , it can profitably deviate by choosing $z - b = a$ and setting a tax $\tau_{Bs} = \tau_{As}^* - \varepsilon$, for all firms in the interval $[0, \hat{s})$. (Part of the additional surplus $z^2/4$ is transferred to the firms in this interval when government B sets $\tau_{Bs} = \tau_{As}^* - \varepsilon$, inducing them to move to B .) This deviation contradicts equilibrium. As there is always an incentive to deviate from $a \neq a^*, b \neq b^*$ no equilibrium can exist. \square

Figure 1



Marginal firm locating in country A

Figure 2

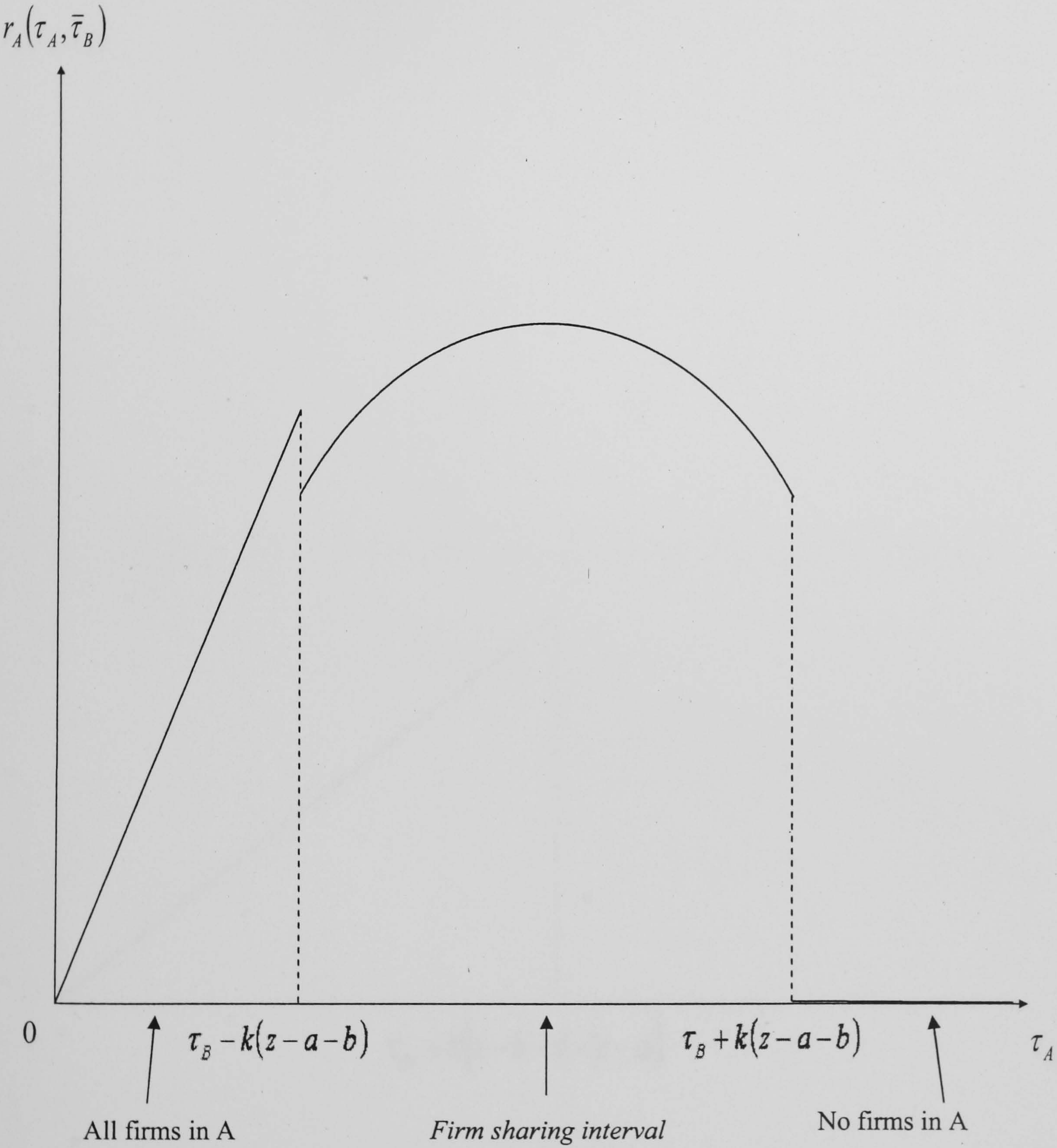
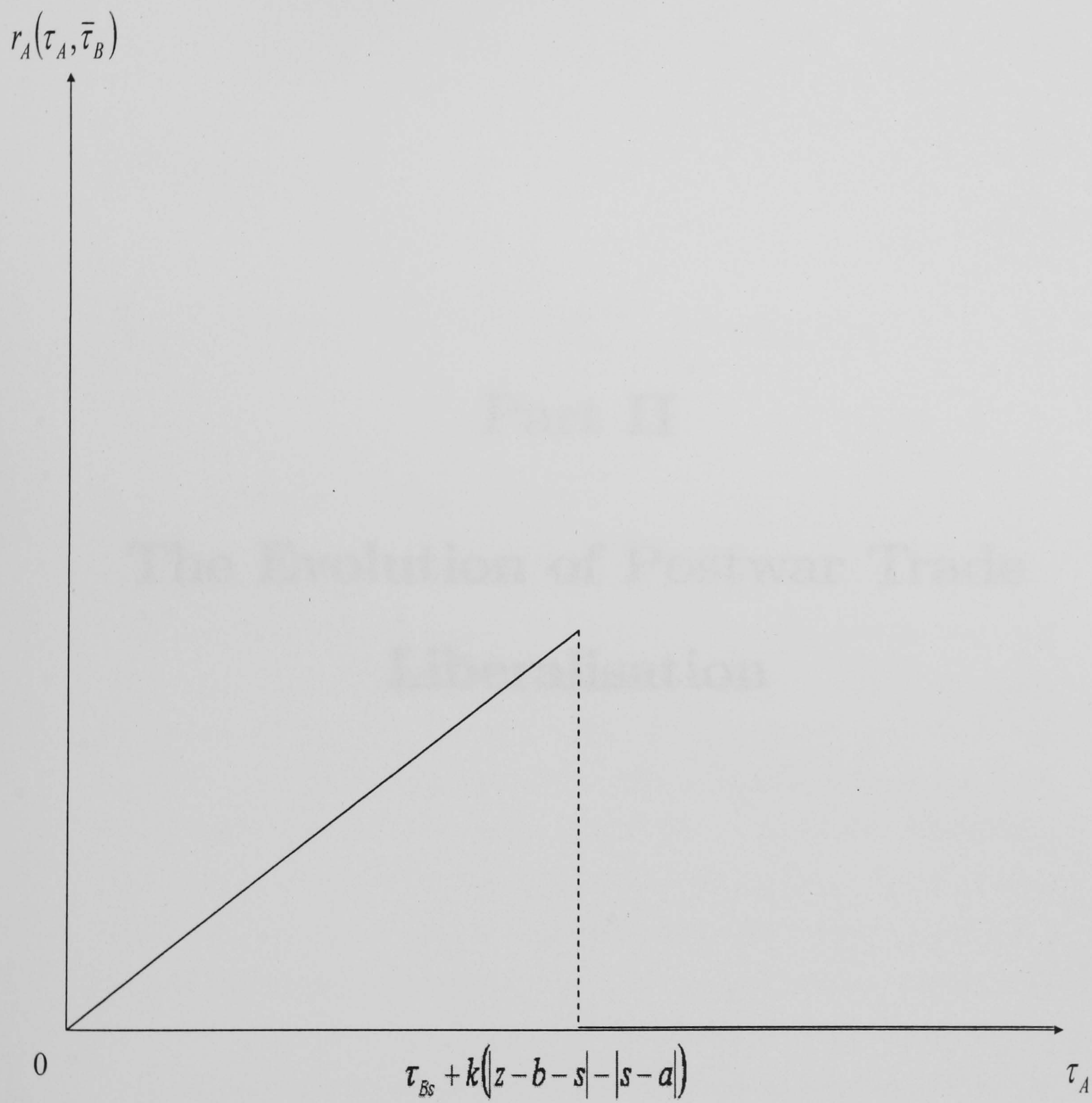


Figure 3



Part II

The Evolution of Postwar Trade Liberalisation

Chapter 3

A Theory of GATT Rounds

3.1 Introduction

The experience of trade liberalization in the period since World War II has presented economists with two puzzles. First, even in developed countries, free trade has remained stubbornly elusive, with average trade-weighted tariffs remaining at low but still positive levels.¹ Second, tariffs have been cut only gradually in successive rounds of negotiations under the General Agreement of Tariffs and Trade (GATT). Since the GATT was drawn up after the war, tariffs have fallen from a trade weighted average of 50 percent to around 5 percent today. Neither of these two facts sits well with the simple textbook view that sees a trade agreement as a simple repeated Prisoner's Dilemma: that is, as a situation where it is individually rational for countries to impose tariffs, but collectively rational to abolish them.

The purpose of this present chapter is to propose an explanation of these two puzzles by modelling the rules imposed on trade liberalization by the GATT. In

¹It could be argued that there is no puzzle in the failure to reach free trade. In a world where trade carries externalities, current positive tariff levels could be efficient. However, in practice there appears to be a consensus that efficiency has not been reached; that mutual gains from trade are still available from further multilateral trade liberalization. To keep things simple, free trade will be used as a metaphor for this 'yet to be obtained' efficient level of international trade.

particular, we will focus on the implications for the liberalization process of the rule on a *withdrawal of equivalent concessions* (WEC) as set out in Article XXVIII, and *tariff bindings* as set out in Article II of the GATT charter. Suppose that a deviant country fails to implement some agreed market access measure, whilst all other parties to the agreement proceed to do so. When the deviation is discovered, under GATT rules trade partners are allowed to do no more than to withdraw market access concessions equivalent to those that the deviant failed to implement. We model exactly this penalty structure in the context of a dynamic game and examine its implications for trade liberalization under the GATT. In terms of the applied game theory literature, WEC imposes *partial irreversibility* on punishments in this game. This is new, in that only *complete irreversibility* has been analyzed in the past (Lockwood and Thomas 2002).

The focus of this chapter is on the broad sweep of trade liberalization under the GATT in the post war period, up to the conclusion of the Uruguay Round in 1996. The idea is to take as read the rules imposed by the GATT regime, and analyze the dynamic equilibrium (liberalization) path that results when the tariff reduction game is played according to these rules. This seems to be a reasonable approach given that the GATT rules were adhered to very closely over that period and did not appear to be questioned. For example, violations of tariff bindings were almost never observed; see Chapter 2 of Whalley and Hamilton (1996) for further details.²

The first main result is that the WEC rule does facilitate trade liberalization but, when retaliation is limited by the WEC rule, free trade certainly cannot be

²In 1996, as part of the conclusion to the Uruguay Round, signatories to the GATT formed the World Trade Organization (WTO). To some extent the analysis of the present chapter is relevant for the period since 1996 too, because the GATT Articles were adopted in the Charter of the WTO (GATT 1994). But since the WTO's inception, we appear to be observing a change in members' attitudes towards the rules of the regime, with a number of instances where rules have been broken. The reasons for this change present an interesting agenda for future research, but will not be taken up here in this present chapter

reached no matter how little countries discount the future. This result contrasts markedly with conventional insights from the theory of repeated games, which indicate that free trade can be achieved, given sufficiently little discounting. The intuition behind this first main result is simple. A standard repeated game allows trade partners to implement the worst (credible) punishment against a deviant. In general, the WEC rule makes such severe punishments illegal. By outlawing a class of severe punishments, the WEC rule compromises efficiency. Note that for this first result, partial irreversibility is imposed only on one side of the agreement. That is to say, WEC limits only the actions of punishers.

The second main result concerns the gradualism of trade liberalization. Specifically, if punishments are constrained by the WEC rule *and* the initial deviation by any country is also constrained, then the most efficient self enforcing path of trade liberalization is gradual. Article 2 of the GATT specifies that a schedule of commitments be maintained. Results of tariff negotiations are recorded as scheduled commitments in the form of tariff bindings; a permanent and irrevocable commitment that tariffs will *not rise above bound levels* for the product in question. This implies that the worst possible deviation is simply not to cut tariffs from the previous period's level (but not to raise them either).³ Under Article 2, because punishment is limited, current tariff cuts can only be made self enforcing by the promise of future tariff reductions. Moreover, if deviation can at worst entail not raising tariffs, then it is always possible to promise liberalization over a number of future periods that would more than compensate. So on the equilibrium path, trade liberalization must take place over a number of periods or 'rounds'.

The chapter builds on a substantial literature going back to Johnson (1953-54).⁴

³In Lockwood and Zissimos (2002) attention is given to the subgame perfection of deviation under Article 2.

⁴Horwell (1966), and more recently Lockwood and Wong (2000) compare trade wars with specific and ad valorem tariffs, showing the outcomes to be different under the respective instruments. Hamilton and Whalley (1983) broaden considerably the basis on which tariff wars can be examined by showing how they can be studied using numerical simulations.

Early contributions explain trade liberalization in a standard repeated game framework, where tariff cuts from their one-shot Nash equilibrium values are explained as the outcome of self enforcing trigger strategies (Dixit 1987).⁵ As remarked above, this trigger strategy approach has two limitations. It cannot explain gradualism and moreover, free trade is always a self-enforcing outcome with sufficiently little discounting. More recent literature has offered several explanations as to why self-enforcing tariff agreements are gradual. The general idea is that initially, full liberalization cannot be self-enforcing, as the benefits of deviating from free trade are too great to be dominated by any credible punishment. But if there is partial liberalization, structural economic change within the domestic economy reduces the benefits of deviation from further trade liberalization (and/or raises the costs of punishment to the deviator). The individual papers differ in their description of the structural change induced by partial liberalization. Staiger (1995) endows workers in the import competing sector with specific skills, making them more productive there than elsewhere in the economy. When they move out of this sector, they lose their skills with some probability. In Devereux (1997), there is dynamic learning-by-doing in the export sector. In Furusawa and Lai (1999), there are linear⁶ adjustment costs incurred when labor moves between sectors. Bond and Park (2000) consider gradualism in a framework where countries are asymmetric.

Finally, Bagwell and Staiger's (1999) work relates closely to our own, in that they too model specifically the GATT/WTO institutional framework. However, their focus is different. First, they make the very important point that the *only* thing that matters in a trade agreement is the terms of trade externality. This

⁵ Among many others, some contributions to the literature on trade agreements that use the threat of retaliation as threat points in cooperative or non-cooperative models include Mayer (1981), Bagwell and Staiger (1990), Bond and Syropoulos (1996) and McLaren (1997). Syropoulos (2001) examines the effect of country size, showing that if one trade partner is larger than another by a significantly large ratio, then it will prefer a trade war to a free trade agreement.

⁶ Furusawa and Lai have an Appendix where they show that with strictly convex adjustment costs, a social planner would choose gradual tariff reduction.

point is made very forcefully by constructing a model that is broader than ours in that it allows for a wider set of political variables to be present. And wider aspects of the GATT institutional framework than just the withdrawal of equivalent concessions are also examined in their work. But our model of withdrawal of equivalent concessions is built around a dynamic game, which theirs is not, and this enables us to bring out some implications of the institutional framework that they do not.⁷ The theory of repeated games has also been used by Bond, Syropoulos and Winters (2001) to study trade block formation, where a *preferential* trade agreement is supported by the credible threat of punishment.

This present chapter also makes a wider contribution to the applied game theory literature on gradualism. In particular, Lockwood and Thomas (2002) study the effect of *complete* irreversibility, showing that irreversibility on the side both of the initial deviant and the punisher are sufficient for gradualism. As pointed out above, in the first part of this present chapter, *partial* irreversibility of the strategic instrument is assumed - here tariffs - on the side of the punisher, but with the initial deviation itself unrestricted. We then see explicitly that gradualism cannot result. Only when there is a degree of irreversibility on both sides does gradualism arise. In this sense, the present chapter extends Lockwood and Thomas (2002).

The chapter proceeds as follows. The next section sets up the basic analytical framework, defines formally the tariff reduction game and a withdrawal of equivalent concessions. Section 3 then defines symmetric equilibrium tariff paths and examines their properties under a withdrawal of equivalent concessions. It is here that we will see how trade liberalization is achieved in this framework but that free trade cannot be reached. Section 4 then examines the circumstances under which gradual trade liberalization can take place, presenting computed equilibrium tariff reduction paths for various parameterizations of a quasi-linear example. Section 5 concludes.

⁷The differences between Bagwell and Staiger's analysis and ours are discussed further in the Conclusions.

3.2 Optimal Tariffs, Trade Agreements and Limited Punishments

3.2.1 Tariffs and Welfare

We work with a simple and standard model of international trade. There are n countries $i \in N$ and the same number of goods. Each country i has an endowment (normalized to unity) of good i (or is endowed with a factor of production that can produce 1 unit of good i). We denote by x_j^i the consumption of good j in country i . The preferences of the representative consumer in country i over $\mathbf{x}^i = (x_j^i)_{j \in N}$ are then⁸

$$u^i(\mathbf{x}^i) = u(x_i^i, \varphi(\mathbf{x}^{-i})) \quad (3.1)$$

where $\mathbf{x}^{-i} = (x_1^i, \dots, x_{i-1}^i, x_{i+1}^i, \dots, x_n^i)$. Also, we assume that in equilibrium, some quantity of imported goods will be consumed i.e. we make the Inada-type assumption that $\lim_{x_j^i \rightarrow 0} \partial u(x_i^i, \varphi(\mathbf{x}^{-i})) / \partial x_j^i = +\infty$, $j \neq i$. An example of this form is the quasi-linear utility function:

$$u^i = x_i^i + \frac{\sigma}{\sigma - 1} \sum_{j \neq i} (x_j^i)^{\frac{\sigma-1}{\sigma}}, \quad i = 1, \dots, n \quad (3.2)$$

with $\sigma > 1$, and where σ measures the elasticity of substitution between different “varieties” of imported goods.

The consumer in country i faces a budget constraint

$$\sum_{j=1}^n p_j (1 + \tau_j^i) x_j^i = p_i + R_i \quad (3.3)$$

where p_j , τ_j^i , R_i are respectively: the world price of good j , the tariff set by country i on good j , and tariff revenue in country i which, as is usually assumed, is returned to the consumer in a lump-sum. Without loss of generality, we set $\tau_i^i = 0$; also note that $-1 < \tau_j^i < \infty$.

⁸We adopt the usual convention that bold characters denote vectors, and non-bold characters denote scalars.

Within a period, $t = 1, 2, \dots$, the order of events is as follows. First, each country i simultaneously chooses an import tariff vector $\tau^i = (\tau_j^i)_{j \in N}$. Then, given world prices $\mathbf{p} = (p_j)_{j \in N}$, and τ^i , the consumer in country $i \in N$ chooses \mathbf{x}^i to maximize u_i subject to the budget constraint, which yields the usual indirect utility function $v^i = v^i(\mathbf{p}, \tau^i, R_i)$ and excess demands. Then, conditional on $\tau = (\tau^1, \dots, \tau^n)$, markets clear and world prices \mathbf{p} for the goods are determined.⁹ These world prices will of course depend on tariffs i.e. $\mathbf{p} = \mathbf{p}(\tau)$, and so will tariff revenues i.e. $R_i = \sum_{j=1}^n p_j(\tau) \tau_j^i x_j^i(\mathbf{p}(\tau))$. We assume that equilibrium prices are unique, given tariffs, so the mapping $\mathbf{p}(\cdot)$ is one-to-one. It is also assumed that technology (embodied in u^i or v^i) is identical across countries.

So, we can write equilibrium welfare of country i , v^i , as a function of $\tau = (\tau^1, \dots, \tau^n)$ only i.e. $v^i = v^i(\tau^1, \dots, \tau^n) \equiv v^i(\mathbf{p}(\tau), \tau^i, R_i(\tau))$. Now we can define a *Nash equilibrium in tariffs* in the usual way as a $\hat{\tau}$ such that $v^i(\hat{\tau}^i, \hat{\tau}^{-i}) \geq v^i(\tau^i, \hat{\tau}^{-i})$, all $\tau^i \in (-1, \infty)^n$, all $i \in N$. We will focus on Nash equilibria where (i) all countries set *common tariffs* i.e. $\hat{\tau}_j^i = \hat{\tau}^i$, all $i \in N$; (ii) all these common tariffs are *equal* $\hat{\tau}^i = \hat{\tau}$, all $i \in N$. Such equilibria exist for the special cases that we consider below, due to the symmetry of the model¹⁰.

We are interested in how fast countries can reduce tariffs from this non-cooperative Nash equilibrium, and also whether they can ever reach free trade i.e. $\tau_j^i = 0$, if the tariff reduction plan must be *self-enforcing* i.e. the outcome of a subgame-perfect equilibrium. It is convenient to impose the constraint that the cooperative tariff reductions have the same structure as does the Nash equilibrium i.e. each country sets a common tariff, τ^i . In this case, we may write country welfare as a function of common tariffs only i.e. $v^i = v^i(\tau^i, \tau^{-i})$. The following result establishes that,

⁹As this is a general equilibrium model, prices are determined only up to a scalar, and so some normalization (e.g. choice of numeraire) must be made. This technical detail, and others, are dealt with in Section 3 below.

¹⁰More generally, it is possible to show that if all $j \neq i$ set the same common tariff, the unique best response of i is to set the same tariff on imports on all countries i.e. a common tariff.

furthermore, countries have symmetric preferences over (common) tariffs.¹¹

Proposition 1. $v^i = v(\tau^i, \tau^{-i})$, and if $\pi(\tau^{-i})$ is any permutation of τ^{-i} , then $v(\tau^i, \tau^{-i}) \equiv v(\tau^i, \pi(\tau^{-i}))$.

For example, if $n = 3$, then $v^1 = v(\tau^1, \tau^2, \tau^3)$, $v^2 = v(\tau^2, \tau^1, \tau^3)$, $v^3 = v(\tau^3, \tau^1, \tau^2)$, and $v(\tau^1, \tau^2, \tau^3) = v(\tau^1, \tau^3, \tau^2)$ etc. We can now use the function v (or, more precisely, functions based on it) to formulate the tariff reduction game precisely. As we are focussing on tariff reductions, we will assume throughout that $\tau = (\tau^1, \dots, \tau^n) \in [0, \hat{\tau}]^n = F^n$.

From now on, for all $\tau, \tau' \in \mathcal{R}_+$, let $w(\tau, \tau') \equiv v(\tau, \tau', \dots, \tau')$ so $w(\tau, \tau')$ is any country i 's payoff in the event that i sets τ , and all $j \neq i$ set τ' . Without much loss of generality, we will assume that w is twice continuously differentiable i.e. let w_1, w_2 be the first partial derivatives of w with respect to τ, τ' respectively. We assume three properties of w :

A1. $w_1(\tau, \tau') \geq 0$, $w_2(\tau, \tau') \leq 0$, for all $(\tau, \tau') \in F^2$, and $w_1(\tau, \tau') > 0$ if $\tau < \hat{\tau}$, $w_2(\tau, \tau') < 0$ if $0 < \tau'$.

A1 asserts that whenever other countries' tariffs are below Nash equilibrium, any country likes an increase in its own (common) tariff, and a reduction in the tariffs of the other countries. In other words, the static tariff game has a Prisoner's Dilemma structure. Our second assumption is very weak:

A2. $w_1(\tau, \tau) + w_2(\tau, \tau) < 0$ for all $(\tau, \tau) \in F^2$ with $\tau > 0$.

This says that any equal reduction in all tariffs, starting from a situation of equal tariffs at or below the Nash level, makes any country better off. Moreover, note that from the optimality of free trade, $w_1(0, 0) + w_2(0, 0) = 0$. Our third assumption is:

A3. $w_{12}(\tau, \tau') < 0$, all $(\tau, \tau') \in F^2$.

¹¹This result, and all others, are proved in the Appendix, where a proof is required.

That is, tariffs are strategic substitutes; the closer other countries' tariffs are to Nash equilibrium tariffs, the smaller the gain any country makes from increasing its own tariff.

Payoffs over the infinite horizon are discounted by a common discount factor δ , $0 < \delta < 1$ i.e.

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^t w(\tau_t^i, \tau_t^{-i}) \quad (3.4)$$

A *game history* at time t is defined as a complete description of past tariffs $h_t = \{(\tau_l^1, \dots, \tau_l^n)\}_{l=1}^{t-1}$. All countries can observe game histories. A *tariff strategy* for country $i = 1, \dots, n$ is defined as a choice of tariffs τ_t^i in periods $t = 1, 2, \dots$ conditional on every possible game history. A *tariff path* of the game is a sequence $\{(\tau_t^1, \dots, \tau_t^n)\}_{t=1}^{\infty}$ that is generated by the tariff reduction strategies of all countries.

Given the symmetry of the model, we restrict our attention to *symmetric* equilibrium¹² tariff paths where $\tau_t^i = \tau_t$, $t = 1, 2, \dots$, i.e. where all countries choose the same tariff in every time period, and we denote such paths by the sequence $\{\tau_t\}_{t=1}^{\infty}$.

3.2.2 Limited Punishments; Withdrawal of Equivalent Concessions

Suppose that $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ is a candidate for an equilibrium tariff sequence, where $\tilde{\tau}_t$ is the tariff “agreed” for period t . Note that there are two kinds of punishment that $i \neq j$ could levy on j for deviating from $\{\tilde{\tau}_t\}_{t=1}^{\infty}$. One is to raise tariffs to the Nash level $\hat{\tau}$, the most severe credible punishment (which we call an *unconstrained* punishment). The other type of punishment is where $i \neq j$, upon observing that j has deviated at time $t-1$, withdraw precisely the equivalent concessions to market access at time t . That is, if the deviant j has set $\tau_{t-1}^j = \tau' > \tilde{\tau}_{t-1}$, then in the next period instead of retaliating by setting $\hat{\tau}$ the other parties withdraw the concessions

¹²In the sequel, it is understood that “equilibrium” refers to subgame-perfect Nash equilibrium.

made, implementing $\tau_t^i = \tau' = \tau_{t-1}^j$ as well. We call this form of punishment payoff a *withdrawal of equivalent concessions* (WEC). In practice, GATT signatories were bound by Article XXVIII to adopt exactly this penalty structure. To support WEC as a subgame perfect equilibrium punishment strategy, there *must* exist an implicit cost to a country of breaking the WEC (i.e. by setting some $\tau_t^i > \tau_{t-1}^j > \tilde{\tau}_{t-1}$ in retaliation). Otherwise, it would never be observed to hold in practice. We denote this cost by c_i . Thus we have a stylized characterization of the GATT rule on withdrawal of equivalent concessions.¹³ Finally, we assume that c_i is so high that no country would wish to violate the WEC rule. Given this, it is clear that the worst credible punishment that the set of countries $N/\{j\}$ can impose on j is to match the deviator's tariff in all subsequent periods.

3.3 Symmetric Equilibrium Paths

3.3.1 Optimal Deviations

We begin by characterizing the optimal deviation from a symmetric equilibrium path $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ for any country i , given that it rationally anticipates that it will be punished by the WEC rule. That is, all $j \neq i$ will match i 's deviation tariff in all subsequent periods if and only if i deviates by setting a tariff $\tau' > \tilde{\tau}_t$. Let i 's optimal deviation at t from the reference path $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ be denoted z_t .

Note that the withdrawal of equivalent concessions applies only to deviation by setting a tariff *above* the agreed rate $\tilde{\tau}_t$. Thus there is an asymmetry in the penalty. Formally, the payoff any country can expect from a deviation to z_t is:

¹³Elsewhere in the literature, reputation effects are modelled explicitly (e.g. Kreps, Milgrom, Roberts and Wilson 1982, Kreps and Wilson 1982, Milgrom and Roberts 1982). Here they are simply introduced by assumption as an enforcement device because we want to focus on the effect of the WEC penal code itself.

$$\Delta(z_t, \{\tilde{\tau}_t\}_{t=1}^{\infty}) = \begin{cases} (1 - \delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t) & \text{if } z_t > \tilde{\tau}_t \\ (1 - \delta)w(z_t, \tilde{\tau}_t) + (1 - \delta) \sum_{s=t+1}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t) & \text{if } z_t < \tilde{\tau}_t \end{cases} \quad (3.5)$$

We are interested in the *optimal deviation* z_t i.e. the choice of z_t that maximizes $\Delta(z_t, \{\tilde{\tau}_t\}_{t=1}^{\infty})$ given the reference path. Due to the discontinuous nature of the payoff $\Delta(z_t, \{\tilde{\tau}_t\}_{t=1}^{\infty})$, an optimal deviation does not exist, but we can precisely bound the gain from deviation. Technically, the largest possible gain from deviation is the supremum of $\Delta(z_t, \{\tilde{\tau}_t\}_{t=1}^{\infty})$ across all values of $z_t \neq \tilde{\tau}_t$, which we denote by $\bar{\Delta}(\{\tilde{\tau}_t\}_{t=1}^{\infty})$.

Lemma 1. Assume A1-A2. Then,

$$\bar{\Delta}(\{\tilde{\tau}_t\}_{t=1}^{\infty}) = \max \left\{ \max_{z_t \geq \tilde{\tau}_t} [(1 - \delta)w(z_t, \tilde{\tau}_t) + \delta w(z_t, z_t)], (1 - \delta) \sum_{s=t}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t) \right\}.$$

This result says that the best that a country can do is either to replicate the payoff on the equilibrium path - the second term in curly brackets - or to deviate by setting tariffs above the agreed level; $z_t \geq \tilde{\tau}_t$. It can never benefit by a unilateral deviation $z_t < \tilde{\tau}_t$.¹⁴ Now, from the first term in curly brackets which gives the gains to deviation, define

$$z(\tau_t) = \arg \max_{z_t \geq \tau_t} \{(1 - \delta)w(z_t, \tau_t) + \delta w(z_t, z_t)\}. \quad (3.6)$$

$z(\cdot)$ can be thought of as a kind of “reaction function” indicating how the optimal deviation varies with the agreed tariff τ_t . We can now obtain a characterization of $z(\cdot)$ that is very useful. Define

$$\zeta(\tau) = \arg \max_z \{(1 - \delta)w(z, \tau) + \delta w(z, z)\} \quad (3.7)$$

¹⁴To see why, recall that a withdrawal of equivalent concessions applies only to upward deviations. If a country were to deviate by setting a tariff that were lower than agreed - $z_t < \tau_t$ - the WEC rule would not require all other countries to follow the deviant downwards. We can therefore ignore the possibility that $z_t < \tau_t$ because, by A1, a country would make itself worse off by deviating in this way.

This is the solution to the problem in equation (3.6), ignoring the inequality constraint. We can think of $\zeta(\tau)$ as a kind of reaction function. Note that

$$\zeta'(\tau) = \frac{(1-\delta)w_{12}(z, \tau)}{D}$$

where $D > 0$ from the second-order condition for the choice of z in (3.7). So, note that if A3 holds, $\zeta'(\tau) < 0$. Also, define $\bar{\tau}$ to satisfy:

$$\bar{\tau} = \zeta(\bar{\tau}) \tag{3.8}$$

This is a self-enforcing tariff level: i.e. at $\bar{\tau}$ the optimal deviation is in fact not to deviate at all.

We now have the following characterization of $z(\cdot)$:

Lemma 2. *Assume A1-A3. Then, there is a unique solution to (3.8), for which $\bar{\tau} < \hat{\tau}$. The solution to (3.6) satisfies: (i) for all $\tau < \bar{\tau}$, $z(\tau) = \zeta(\tau) \geq \bar{\tau} > \tau$; (ii) for all $\tau \geq \bar{\tau}$, $z(\tau) = \tau$.*

We now have a complete characterization of the optimal deviation z_t , given any tariff τ_t . So for any $\tilde{\tau}_t$ in a candidate equilibrium sequence $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ we know the optimal deviation for that period under WEC. This will now be used to characterize uniquely the efficient equilibrium path.

3.3.2 Efficient Equilibrium Paths and Failure to Reach Free Trade

We can now formally define the conditions that must hold if a symmetric tariff path is to be a subgame-perfect one in our game. In every period, the continuation payoff from the path must be at least as great as the maximal payoff from deviation, given that a punishment consistent with the WEC will ensue. From Lemma 1, the maximal relevant payoff from deviation at t is $(1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t)))$. So, formally, we require:

$$(1-\delta)(w(\tau_t, \tau_t) + \delta w(\tau_{t+1}, \tau_{t+1}) + \dots) \geq (1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t))), \quad t = 1, \dots$$

(3.9)

Of course, a whole set of paths will satisfy this sequence of inequalities: let this set of equilibrium paths be denoted E . An *efficient tariff reduction path* in the set E is simply a sequence $\{\tau_t\}_{t=1}^{\infty}$ of tariffs in E for which there is no other sequence $\{\tau'_t\}_{t=1}^{\infty}$ also in E which gives a higher payoff to any country, as calculated by (3.4). Following the arguments of Lockwood and Thomas (2002), it can be shown that if $\{\tau_t\}_{t=1}^{\infty}$ is efficient, (3.9) holds with equality at every date i.e. :

$$(1-\delta)(w(\tau_t, \tau_t) + \delta w(\tau_{t+1}, \tau_{t+1}) + \dots) = (1-\delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), z(\tau_t)), \quad t = 1, \dots \quad (3.10)$$

The intuition is that if (3.9) held with strict inequality, it would be possible to reduce the tariff path by a small amount without violating (3.9).

Of the class of equilibrium paths E , it is obviously the efficient path (shown to be unique below) that is of most interest. We now turn to characterizations of the efficient equilibrium path. Our first main result, Proposition 2, establishes that free trade is in fact impossible under WEC.

Proposition 2. (*Failure to reach free trade*) *Let $\{\tau_t\}_{t=1}^{\infty}$ be an equilibrium path. Then $\tau_t > 0$, for all $\delta < 1$, all t .*

The proof of this Proposition works by showing that if all other countries agree to adopt free trade at *any* point in time, then the last will have an incentive to deviate by levying a positive tariff. So such an agreement would not be self-enforcing. This is clearly in contrast to the standard case with unlimited punishments. For in that case, countries can credibly punish deviators by reverting to (for example) Nash tariffs, and then it is well-known that for some $\delta_0 < 1$, free trade can be attained in equilibrium for all $\delta > \delta_0$. Instead, Proposition 3 is reminiscent of the results of Lockwood and Thomas(2002), who study a repeated prisoner's dilemma with *complete* irreversibility of actions.

We now turn to the more difficult question of what form the efficient path takes. Say that an equilibrium tariff reduction path is a *stationary* path if $\tau_t = \tau$, all $t \geq 1$ (recall $\tau_0 = \hat{\tau}$); that is, there is an immediate and permanent tariff reduction. A stationary equilibrium path must satisfy:

$$\alpha(\tau) \equiv \max_{z \geq \tau} \{(1 - \delta)w(z, \tau) + \delta w(z, z)\} \leq w(\tau, \tau) \equiv \beta(\tau).$$

To characterize such paths, note first the properties of α, β . First, β is decreasing in τ by A2, and α is decreasing by A1, A2. Second, at the Nash equilibrium, as $z = \hat{\tau}$ is a best response to $\hat{\tau}$, $\alpha(\hat{\tau}) = \beta(\hat{\tau})$ i.e. the non-cooperative Nash equilibrium is a stationary equilibrium path¹⁵. Third,

$$\alpha(0) \equiv \max_{z \geq 0} \{(1 - \delta)w(z, 0) + \delta w(z, z)\} > w(0, 0) \equiv \beta(0)$$

as a small increase in z from 0 strictly increases $w(z, 0)$ (from A1), while leaving $w(z, z)$ unchanged (as $w(z, z)$ is maximized at zero, by A2).

So, the possibilities are shown in Figure 1. Next, as α, β are both downward-sloping, they may have multiple crossing-points, as shown. Note that $\alpha(\tau)$ and $\beta(\tau)$ coincide over the range $\bar{\tau} \leq \tau \leq \hat{\tau}$. This is because, by Lemma 2, $z(\tau) = \tau$ for all $\tau \geq \bar{\tau}$. So

$$\begin{aligned} \alpha(\tau) &= \max_{z \geq \tau} \{(1 - \delta)\psi(z, \tau) + \delta\psi(z, z)\} \\ &= \psi(\tau, \tau) = \beta(\tau) \text{ for all } \tau \geq \bar{\tau} \end{aligned}$$

Finally, the *smallest* stationary equilibrium tariff will be at the lowest crossing point of α, β , namely τ^* . Moreover, using Lemma 2, it is possible to show that under some additional assumptions, $\tau^* = \bar{\tau}$. Formally, we have:

Proposition 3. *Let $\tau_0 = \hat{\tau}$. There is a unique efficient stationary path, $\tau_t = \tau^*$, all $t \geq 1$, where $\tau^* > 0$ is the smallest root of the equation $\alpha(\tau) = \beta(\tau)$. Moreover, if A3 holds, and $w_{11}(\tau, \tau), w_{22}(\tau, \tau) \leq 0$ on $[0, \tau]$, then $\tau^* = \bar{\tau} < \hat{\tau}$.*

¹⁵Note that it is *not* claimed that $\hat{\tau} = \zeta(\hat{\tau})$. In fact, it is easily checked from the definition of (3.7) that $\zeta(\hat{\tau}) < \hat{\tau}$, so the constraint $z \geq \hat{\tau}$ in the definition of α binds, implying that $z(\hat{\tau}) = \hat{\tau}$, and consequently, that $\alpha(\hat{\tau}) = (1 - \delta)\psi(\hat{\tau}, \hat{\tau}) + \delta\psi(\hat{\tau}, \hat{\tau}) = \psi(\hat{\tau}, \hat{\tau}) = \beta(\hat{\tau})$.

Proposition 3 shows that under a withdrawal of equivalent concessions it is possible for all countries to agree to reduce tariffs immediately to the level $\bar{\tau}$, holding them there indefinitely, and moreover, this is the best equilibrium stationary path. The result is illustrated in Figure 2, which refines Figure 1.

The question then arises as to whether there is a *non-stationary* path in E which is more efficient than the stationary path $\tau_t = \bar{\tau}$, $t \geq 1$. The following result answers this negatively:

Proposition 4. *The stationary path, which has $\tilde{\tau}_t = \bar{\tau}$, all $t \geq 1$, is the unique efficient path in E .*

The idea of the proof is the following. If there is a more efficient equilibrium path, then it must involve a tariff $\tau_t < \bar{\tau}$. But, the dynamics of (3.10), expressed as a difference equation, tell us that once $\tau_t < \bar{\tau}$, $\tau_{t+1} < \tau_t$ i.e. the path must be monotonically decreasing. But this is impossible, as either it implies a stationary equilibrium path below $\bar{\tau}$ (impossible by definition), or a tariff sequence diverging to minus infinity (which cannot be efficient).

We now illustrate our results with the quasi-linear example i.e. we assume that preferences take the form (3.2). This example is analyzed thoroughly in the appendix. First, it can be shown that the Nash equilibrium tariff is $\hat{\tau} = 1/(\sigma - 1)$. Also, we show that

$$\bar{\tau} = \frac{1 - \delta}{\sigma(1 + \delta) - 1}. \quad (3.11)$$

Note from (3.11) that in general, $0 < \bar{\tau} < \hat{\tau}$. That is, $\bar{\tau} \rightarrow \hat{\tau}$ as $\delta \rightarrow 0$, and $\bar{\tau} \rightarrow 0$ as $\delta \rightarrow 1$. When agents place a high weight on future outcomes, tariff rates close to zero can be achieved under WEC. The elasticity of substitution between goods is also inversely related to the level of $\bar{\tau}$.

If the GATT provides a means by which countries select the efficient tariff reduction path, then Propositions 2, 3 and 4 provide a complete characterization of this path. Accordingly, under WEC trade liberalization can be achieved, but free trade cannot be reached. However, at present our model cannot “explain” the

gradualism in tariff-cutting observed in practice.

3.4 Loss of Political Good Will and Gradual Tariff Reduction

In Section 3.2.2, it was argued that there must exist an implicit cost to countries of breaking the WEC penal code. If not, then it would never actually be observed to hold. This cost was posited as a loss of political good will, which would be exerted in other areas of the international political arena. This loss of political good will is now extended to the initial deviant. Specifically, we will assume the following. If country i sets $\tau_t > \tau_{t-1}$, it incurs a *political* cost of deviation \tilde{c}_i .¹⁶ If on the other hand $\tau_t \leq \tau_{t-1}$, country i incurs no such cost at the initial deviation.¹⁷

A justification of this penalty structure is as follows. Article 2 of the GATT specifies that a schedule of commitments be maintained. Results of tariff negotiations are dutifully recorded as scheduled commitments in the form of tariff bindings; a permanent and irrevocable commitment that tariffs *will not rise above bound levels* for the product in question. Violations of tariff bindings become the subject of dispute settlement; with initial complaint, investigation and hearing before panels, panel findings, and rulings by the GATT council to come into compliance. Failure to return to compliance will eventually lead to retaliation being sanctioned by the GATT on the part of parties affected by the violation of bindings against violators.

Why has this been so? Why have tariff bindings under GATT de facto become permanent and irreversible commitments, and what has been the penalty struc-

¹⁶We do not assume in general that $c_i = \tilde{c}_i$. For example, it may be that the political cost of reneging on the original agreement in the first place is higher than the cost of deviating later, in the punishment phase. Or there may be a higher cost to losing the moral high ground.

¹⁷Note that a country can deviate from the agreement without incurring a loss of political good will by setting τ' so that $\tilde{\tau}_t < \tau' < \tau_{t-1}$.

ture to maintain this system? Firstly, tariff bindings have acquired the status of an international commitment comparable to that of other international treaties. Bindings, if committed to, effectively slot into a box of enshrined cross country commitments comparable to military and diplomatic treaties (Jackson 1989 chapters 2, 4). Violation of tariff bindings brings into question the soundness of a country's financial commitments, its trustworthiness in strategic and military matters, its diplomatic reputation. Violating tariff bindings incurs large costs outside the tariff area (Keohane 1982, 1984 chapter 4).¹⁸

It is somewhat unsatisfactory that these political costs of tariff reversals are not firmly micro-founded. However, it appears that such costs exist and are important in the international arena. And no theory exists of which we are aware to explain the impact on tariff reductions of this type of cost. Therefore, in the absence of such a theory, it seems appropriate to simply assume that such costs exist in order to examine their consequences.

We will assume in what follows that \tilde{c}_i is high enough so that a deviation at t will never be above τ_{t-1} and thus incur loss of political goodwill. We can now reformulate the equilibrium conditions (3.9) under this new constraint. It is clear that in the event that a country deviates, the "optimal" deviation given in (3.6) will be chosen, unless $z(\tau_t) > \tau_{t-1}$, in which case τ_{t-1} will be chosen. So, defining $\chi(\tau_t, \tau_{t-1}) = \min\{z(\tau_t), \tau_{t-1}\}$, the equilibrium conditions become

$$(1 - \delta)(w(\tau_t, \tau_t) + \delta w(\tau_{t+1}, \tau_{t+1}) + \dots) \geq \quad (3.12)$$

$$(1 - \delta)w(\chi(\tau_t, \tau_{t-1}), \tau_t) + \delta w(\chi(\tau_t, \tau_{t-1}), \chi(\tau_t, \tau_{t-1})), \quad t = 1, \dots$$

As before, let the set of equilibrium tariff paths be E , and define the efficient tariff paths in E as those paths that maximize (3.4). Also as before, any efficient path

¹⁸We thank John Whalley for suggesting this synthesis of work by Jackson and Keohane in support of our present argument.

Current (at the time of writing) protectionist measures, imposed on steel imports by the European Union and US, appear to be in breach of tariff bindings. Yet over the postwar period in general, the focus of this chapter, instances of violations of tariff bindings were rare.

must satisfy (3.12) with equality.

To proceed, we first introduce the following result. By Lemma 2, we know that $z(\tau_t) \geq \bar{\tau}$ for all $\tau_t < \bar{\tau}$. So, $z(\tau_t) > \tau_{t-1}$ also if $\tau_{t-1} < \bar{\tau}$. Formally:

Lemma 3. *If $\tau_t, \tau_{t-1} \leq \bar{\tau}$, then $\chi(\tau_t, \tau_{t-1}) = \tau_{t-1}$.*

This says that as long as $\tau_t, \tau_{t-1} \leq \bar{\tau}$, the optimal retaliation is τ_{t-1} . Recall from Lemma 2 that $\zeta(\tilde{\tau}) > \bar{\tau}$ if $\tilde{\tau} < \bar{\tau}$. But now a loss of political good will prohibits a deviation to this level because $\zeta(\tilde{\tau}) = z(\tau_t) > \bar{\tau} > \tau_{t-1}$. If the cost from a loss of political good will is high enough, the country is better off adopting τ_{t-1} rather than $\zeta(\tilde{\tau}) = z(\tau_t)$; ie $\chi(\tau_t, \tau_{t-1}) = \min\{z(\tau_t), \tau_{t-1}\} = \tau_{t-1}$.

Now, suppose that $\{\tau_t\}_{t=s}^{\infty}$ is an efficient path from s onwards with $\tau_t \leq \bar{\tau}$, $t \geq s$. From (3.12) and Lemma 3, this path must satisfy

$$\begin{aligned} (1 - \delta)(w(\tau_t, \tau_t) + \delta w(\tau_{t+1}, \tau_{t+1}) + \dots) &= \\ (1 - \delta)w(\tau_{t-1}, \tau_t) + \delta w(\tau_{t-1}, \tau_{t-1}), &t = 1, \dots \end{aligned} \quad (3.13)$$

Advancing (3.13) one period, multiplying both sides by δ , subtracting from (3.13), and dividing the result by $1 - \delta$, we get:

$$w(\tau_t, \tau_t) = w(\tau_{t-1}, \tau_t) + \frac{\delta}{1 - \delta} w(\tau_{t-1}, \tau_{t-1}) - \delta \left[w(\tau_t, \tau_{t+1}) + \frac{\delta}{1 - \delta} w(\tau_t, \tau_t) \right] \quad (3.14)$$

which is a second-order difference equation¹⁹ in τ_t . This can be seen more clearly by rearranging (3.14) to get:

$$w(\tau_t, \tau_{t+1}) = \frac{1}{\delta} [w(\tau_{t-1}, \tau_t) - w(\tau_t, \tau_t)] + \frac{w(\tau_{t-1}, \tau_{t-1})}{1 - \delta} - \frac{\delta w(\tau_t, \tau_t)}{1 - \delta}, \quad t > 1. \quad (3.15)$$

Let $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ be the sequence that solves (3.15) with initial conditions τ_0, τ_1 . We can now establish gradualism by showing that as long as there is a tariff reduction in the first period then tariffs must strictly fall in all subsequent periods along any efficient equilibrium path.

¹⁹This is an unusual difference equation in that it has a continuum of stationary solutions i.e. setting $\tau_{t-1} = \tau_t = \tau_{t+1}$ always solves (3.4).

Lemma 4. Any sequence $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ that satisfies (3.15), with initial conditions τ_0, τ_1 with $0 < \tau_1 < \tau_0$ is strictly decreasing i.e. $0 < \tau_{t+1}(\tau_0, \tau_1) < \tau_t(\tau_0, \tau_1)$ all $t \geq 1$.

Now consider the construction of an efficient path, given these results. First, τ_0 is given at $\hat{\tau}$. Second, from $t = 2$ onwards, i.e. conditional on τ_0, τ_1 , the unique efficient path is simply $\{\tau_t(\tau_0, \tau_1)\}_{t=2}^{\infty}$ as long as (i) $\tau_1 < \tau_0$ (required by Lemma 4), and (ii) $\tau_1 \leq \bar{\tau}$ (required by Lemma 3: otherwise, the efficient path does not satisfy (3.15)). So, it remains to choose $\tau_1 \leq \bar{\tau} < \hat{\tau}$. If the path is to be efficient, the incentive constraint (3.12) must hold with equality in period 1 i.e.

$$\begin{aligned} & (1 - \delta)(w(\tau_1, \tau_1) + \delta w(\tau_2(\hat{\tau}, \tau_1), \tau_2(\hat{\tau}, \tau_1)) + \dots) \\ &= (1 - \delta)w(\chi(\tau_1, \hat{\tau}), \tau_1) + \delta w(\chi(\tau_1, \hat{\tau}), \chi(\tau_1, \hat{\tau})) \end{aligned} \quad (3.16)$$

We now have:

Proposition 5. There exists a smallest value of τ_1 , $0 < \tilde{\tau}_1 < \bar{\tau}$ that satisfies (3.16). Consequently, the path $(\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3, \dots)$ is the unique efficient path, with $\tilde{\tau}_t = \tau_t(\hat{\tau}, \tilde{\tau}_1)$, $t > 1$. This path exhibits a gradually decreasing tariff i.e. $\tilde{\tau}_{t+1} < \tilde{\tau}_t$, $t \geq 1$.

From Proposition 5 we learn that it is possible to achieve an equilibrium path for which $\tilde{\tau}_t < \bar{\tau}$, all $t \geq 1$. Consider some period s in which tariffs have been reduced by a gradual process over periods $t = 1, \dots, s - 1$ to some tariff level $\tilde{\tau}_s < \bar{\tau}$. Now suppose that the agreement requires $\tilde{\tau}_{s+1} < \tilde{\tau}_s$ in period $s + 1$. If the agreement proposes no further reductions in future periods, then country i may do better by maintaining $\tilde{\tau}_s$ in $s + 1$ whilst all other countries proceed to set $\tilde{\tau}_{s+1} < \tilde{\tau}_s$, even if all countries impose the WEC penal code in all periods after that. But it is always possible to promise additional reductions in future periods that can compensate for the gains to deviation in period s .

Why is the cost from loss of political good will necessary for this process? In its absence, the unilateral gains from deviating to $\bar{\tau}$ are greater than the gains from all future reductions. Indeed, the gains from deviation grow with the size of the

overall reduction. But if a loss of political good will limits a deviation to the tariff level in the previous period, $\tilde{\tau}_{t-1}$, then the promise of all future reductions *can* be large enough to compensate for the gains from deviation in a single period.

Proposition 5 establishes that we can restrict attention to a tariff reduction sequence $\{\tilde{\tau}_t\}_{t=1}^{\infty}$ for which $0 < \tilde{\tau}_t \leq \bar{\tau}$ all $t \geq 1$. We are able to explore the properties of the efficient equilibrium tariff reduction path further by looking at specific examples. To do this, the functional form must be specified and it must be verified that for the example under consideration assumptions A1-A3 are satisfied.

3.4.1 Computed Paths

A computational algorithm is used to find the efficient equilibrium path. This entails finding the smallest possible value of $\tau_1 - \tilde{\tau}_1$ - that satisfies (3.16) and therefore, by Proposition 5, gives rise to the unique efficient path $\tilde{\tau}_t = \tau_t(\hat{\tau}, \tilde{\tau}_1)$. There exists no analytical way of finding $\tilde{\tau}_1$, but it can be approximated in the following way. First, set the second initial condition, τ_1 , of the difference equation defined in (3.16) equal to the efficient stationary tariff, $\bar{\tau}$. (Recall that the first condition is fixed at $\tau_0 = \hat{\tau}$.) Then reduce this second initial condition by a small step ε and check that the resulting difference equation converges to some positive tariff rate. Continue in this way, reducing τ_1 by steps of ε until it is so low that the difference equation diverges. The final convergent difference equation is then the approximation to the efficient path. The approximation is more accurate the smaller the step size ε . Intuitively, the efficient tariff reduction path cannot bring about non-positive tariffs, because free trade cannot be reached, by Proposition 2.

The algorithm is as follows:

1. Let $k = k + 1$.
2. Set $\tau_0 = \hat{\tau}$ and $0 < \tau_1 = \bar{\tau} - k\varepsilon \leq \bar{\tau}$ as initial conditions and solve (3.15) forward for T periods.
3. If $\tau_T(\tau_1; \tau_0, \delta) > 0$, set $S_k = S_{k-1} \cup \{\hat{\tau} - k\varepsilon\}$ and go to 1.

4. If $\tau_T(\tau_1; \tau_0, \delta) \leq 0$, stop. Discard this path.

This algorithm is initialized by setting $S_0 = \emptyset$. Note that the algorithm can only run at most for m steps, where m is the largest integer smaller than $\bar{\tau}/\varepsilon$. Let $K+1 \leq m$ be the number of steps after which the algorithm stops. The algorithm stops when a path fails the criterion of $\tau_T(\tau_1; \tau_0, \delta) > 0$. Having failed, this last path must be discarded. Then $S_K = S_{K-1} \cup \tau_1 - K\varepsilon$ and $\tau_1 = \bar{\tau} - K\varepsilon$ is the smallest member. S_K then comprises the full set of tariff reduction paths that satisfy (3.16), and $\tau_1 = \bar{\tau} - K\varepsilon = \tilde{\tau}_1$ gives rise to the efficient path, as required.

The technical details are as follows. The utility function (3.2) is substituted into the second order difference equation that defines an equilibrium tariff reduction path (3.15). The resulting expression is used to solve sequentially for the equilibrium tariff level τ_{t+1} , given levels in τ_{t-1} and τ_t . Recall that the algorithm requires the size of the steps between simulations ε and the total number of periods T to be determined. We use, respectively, $\varepsilon = 0.0001$ and $T = 10000$. A smaller value of ε and a larger value of T would yield greater accuracy in computation of the equilibrium reduction path, but take longer.

The procedure is begun with $k = 0$, so in calculating S_0 the procedure is initialized using $\tau_0 = \hat{\tau}$, $\tau_1 = \bar{\tau}$. Let K be the highest value of k for which $\tau_T(\tau_1; \tau_0, \delta) > 0$. The algorithm is illustrated in Figure 3, for $\sigma = 2$ and $\delta = 0.5$, where the path corresponding to step $k = K$ is the approximation to the efficient tariff reduction path. The tariff level is shown on the vertical axis, with simulation periods on the horizontal axis. Only the first 1000 periods of the simulation are presented. We also show what happens for $k = K+1$ and $k = K+2$. Note that no value for the number of countries is specified. The reason is that n has no impact whatever on the equilibrium path under the quasi-linear preference specification.²⁰

Given $\sigma = 2$, $\delta = 0.5$, and $\tau_0 = \hat{\tau} = 1$ we have $\tau_1 = \bar{\tau} = 0.25$ for $k = 0$ and $\tau_1 = 0.2499$ for $k = 1$ and so on. One of the paths shown in Figure 3 is for

²⁰To put this another way, if a closed form solution for the reduction path could be found, then n would cancel from the expression.

$k = K = 1426$, so that $\tau_1 = 0.1704$. Note that for this set of initial conditions, the reduction path stabilizes; $\tau_{10000} = 0.102748 > 0$. This is the efficient gradual reduction path. How do we know? When k is increased by 1 to $K + 1 = 1427$, the criterion $\tau_T(\tau_2; \tau_1, \delta) > 0$ fails.

This path that fails the criterion is also presented in Figure 3. Observe that $k = K + 1 = 1427$ implies $\tau_1 = 0.1703$. The path diverges sharply downwards and τ_{10000} - were it to be displayed - would be significantly below 0, failing the criterion for that path to be an equilibrium. At $t = 100$, $\{\tau_{100}(\bar{\tau} - (K + 1)\varepsilon; 1, 0.5)\} = 0.099384$, and is close to $\{\tau_{100}(\bar{\tau} - K\varepsilon; 1, 0.5)\}$. However, as t increases further the path of the sequence $\{\tau_t(\bar{\tau} - (K + 1)\varepsilon; 1, 0.5)\}_{t=1}^T$ diverges downwards sharply from $\{\tau_t(\bar{\tau} - K\varepsilon; 1, 0.5)\}_{t=1}^T$, so $\tau_T(\tau_1; \tau_0, \delta) \leq 0$ for $K + 1$ and the path must be discarded (see Step 4 of the algorithm above). For $K + 2$, where $\tau_1 = 0.1702$, the divergence takes place at an even lower value of t .

Figure 3 also shows the one off tariff reduction path, with the tariff being reduced immediately to $\bar{\tau}$ in period 1. Between this tariff and the most efficient tariff reduction path lies the ‘Region of gradual reduction paths’ which (in the limit) fills the area between the one off reduction path and the efficient gradual reduction path.

On a cautionary note, the algorithm may pick a path that appears to approximate the equilibrium path for a given value of T , but fails for some larger T . In view of this possibility the value of K and corresponding τ_1 for the optimal path given here by $\tau_1 = 0.1704$ was checked for robustness by setting $T = 100000$ and verifying that $\tau_T(\tau_1; \tau_0, \delta) > 0$ continued to hold. The same robustness check was also performed on all other computed optimal paths presented below.

Figures 4 and 5 illustrate efficient tariff reduction paths that result from comparative dynamics exercises carried out using the quasi-linear preference function on the same format as Figure 3. These latter figures present only the first 250 of 10000 periods. Figure 4 shows how the optimal reduction path varies with the substitution elasticity σ , whilst Figure 5 indicates the impact of variation in the

discount factor δ .

Look at Figure 4 first. There are optimal reduction paths for three substitution elasticities $\sigma = 2, 5$ and 10 with the other parameter held fixed at $\delta = 0.5$. The key data and results for these simulations are presented in boxes on the far right hand side of the figure. As in Figure 3, for each value of σ we already know $\hat{\tau}$ and $\bar{\tau}$ from the analysis. Both are decreasing in σ , and the figure shows that the optimal reduction paths are monotonically decreasing in σ as well, as one would expect.

The discount rate δ only affects the reduction path, and not $\hat{\tau}$, explaining why the optimal reduction paths in Figure 5 start at the same point and decline towards different limits. Simulations for $\delta = 0.1, 0.5$ and 0.9 are shown, holding $\sigma = 2$ constant. We see that for higher values of δ the liberalization path exhibits greater liberalization at each point in time t .

3.5 Conclusions

This present chapter helps to explain two stylized facts about trade liberalization, namely failure to reach free trade and gradualism, by studying the interplay between countries' unilateral incentive to set tariffs and the institutional structure set up in the framework of the GATT to achieve trade liberalization, paying special attention to the role of time in the process. We use a dynamic game framework, which makes it possible to take account of the fact that a country is able to renege on an agreement for some time before being found out. In addition, the GATT institutional structure limits the extent of allowable retaliation. It is the interaction of these two features in our model, novel in the present context, which enables us to explain the failure to reach free trade and gradualism.

We return to an apparent difference in the outcome from our modelling framework to that of Bagwell and Staiger (1999). They also model a trade agreement using a penalty structure based on the GATT's withdrawal of equivalent conces-

sions as a penalty structure. However, in their model it is possible to achieve full efficiency whilst in ours it is not. In their conclusion, they point out that there may in fact be enforcement difficulties. (As Bagwell and Staiger point out, enforcement difficulties have been studied in a wider context by Dam 1970). Our dynamic game captures and formalizes an element of this enforcement difficulty that Bagwell and Staiger's model does not; that a country is able to reap the benefits of deviation for a period before retaliation occurs. It is this that drives the inability to obtain full efficiency in our model, which is not a feature of Bagwell and Staiger's.

Inevitably, the theoretical framework simplifies the situation in a number of key respects. All countries are assumed to be symmetrical, and small in terms of their purchasing power on world markets relative to the political costs of raising protectionism. Each country exports only a single good, with all countries equally open at a given time. In practice countries export a number of goods, with levels of openness varying across sectors. Variation in country size and purchasing power across different markets is likely to make the actual dynamics of perpetual liberalization considerably more subtle and complex, with more rapid progress achieved in areas where countries receive greater gains from protectionism relative to the political costs incurred. Gradualism in a context where there are asymmetries across countries has been studied by Bond and Park (2000), but not within the context of the GATT penalty structure that we examine here. By defining a symmetrical modelling framework this issue is completely suppressed in our present chapter.

Another simplification of the present framework is that there are no stochastics in the present model. In the real world, an important aspect of the negotiating rounds under the GATT must be to renegotiate existing arrangements in the light of technological innovations. The proposed agenda for negotiations in the latest round at Doha is dominated by electronic and telecommunications based commerce, both areas that did not exist when the GATT was first drawn up. The framework of the present paper does not allow for such innovations. But it presents

a framework on which an allowance for such innovations might be built.

A promising direction for future research would allow trade block formation to be considered. The theory of repeated games has been used to study trade block formation, where a preferential trade agreement is supported by the credible threat of punishment. In a recent paper using a repeated game framework Bond, Syropoulos and Winters (2001) point out that trade liberalization within the European Union has been very slow. It may be that our framework provides a way of understanding gradualism between members.

There may be many other competing pressures other than the standard terms-of-trade motive working against further liberalization, and these are also suppressed in our model. One area that has attracted significant attention recently is the incentive for politicians to be protectionist in order to gain financial backing from industrialists (Grossman and Helpman 1994, 1995a) and for electorates to elect politicians who signal that they will adopt protectionist measures in order to increase their chances of being elected (Lausel and Riezman 2001). These protectionist forces may be outweighed at an early stage by the gains that we describe which are relatively large early on in the process, but not later once the potential gains become relatively small. Future research could study the interaction of these counteracting forces.

3.6 Appendix

3.6.1 Proof of Propositions

Proof of Proposition 1. Fix $i \in N$, and normalize prices by setting $p_i = 1$, so $\mathbf{p} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n)$. Then, by the symmetry of the model, and taking τ_i as fixed,

$$\mathbf{p}(\tau^i, \pi(\tau^{-i})) = \pi(\mathbf{p}(\tau^i, \tau^{-i})), \quad R^i(\tau^i, \tau^{-i}) = R^i(\tau^i, \pi(\tau^{-i})) \quad (3.17)$$

where $\pi(\cdot)$ is any permutation function i.e. a permutation in tariffs of other countries leads to the same permutation in their equilibrium prices, as tariffs are the only variables affecting excess demands that differ across countries. Now note that by definition,

$$v^i(\tau^i, \tau^{-i}) \equiv v^i(\mathbf{p}(\tau^i, \tau^{-i}), \tau^i, R^i(\tau^i, \tau^{-i})) \quad (3.18)$$

Also, by symmetry of the model,

$$v^i(\pi(\mathbf{p}(\tau^i, \tau^{-i})), \tau^i,) = v^i(\mathbf{p}(\tau^i, \tau^{-i}), \tau^i, R^i) \quad (3.19)$$

i.e. country utility is the same if the world prices of imports are permuted. So we have

$$\begin{aligned} v^i(\tau^i, \pi(\tau^{-i})) &= v^i(\mathbf{p}(\tau^i, \pi(\tau^{-i})), \tau^i, R^i(\tau^i, \pi(\tau^{-i}))) \\ &= v^i(\pi(\mathbf{p}(\tau^i, \tau^{-i})), \tau^i, R^i(\tau^i, \tau^{-i})) \\ &= v^i(\mathbf{p}(\tau^i, \tau^{-i}), \tau^i, R^i(\tau^i, \tau^{-i})) \\ &= v^i(\tau^i, \tau^{-i}) \end{aligned} \quad (3.20)$$

where the first line of (3.20) is from (3.18), the second is from (3.17), the third is from (3.19), and the fourth is from (3.18) again. This proves the second part of the Lemma. To prove the first part, note that as all countries are identical up to a permutation of the indices of the goods, $v^j = v^i(\tau^j, \tau^{-j})$, all i, j so $v^i = v(\tau^i, \pi(\tau^{-i}))$ as required. \square

Proof of Lemma 1. (a) First, suppose that a country deviates to $z_t < \tilde{\tau}_t$. Then, from (3.5), as there is no retaliation, future payoffs are unaffected by the choice of deviation. Moreover, as is increasing in z_t by A1, the payoff to deviation of the form $z_t < \tilde{\tau}_t$ is increasing in z_t . Therefore, there is no optimal deviation, but the supremum of the payoff to this kind of deviation is

$$\lim_{z_t \rightarrow \tilde{\tau}_t} [w(z_t, \tilde{\tau}_t)(1-\delta)w(z_t, \tilde{\tau}_t) + (1-\delta) \sum_{s=t+1}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t)] = (1-\delta) \sum_{s=t}^{\infty} \delta^{s-t} w(\tilde{\tau}_t, \tilde{\tau}_t)$$

(b) If a country deviates to $z_t > \tilde{\tau}_t$, it receives

$$g(z_t, \tau_t) = (1 - \delta)w(z_t, \tau_t) + \delta w(z_t, z_t) \quad (3.21)$$

So, it suffices to show that (3.21) has a global maximum z_t^* on (τ_t, ∞) . If this is *not* the case, then there exists an increasing sequence $\{z^n\}$ with $\lim_{n \rightarrow \infty} z^n \rightarrow \infty$, for which $g(z^n, \tau_t)$ is monotonically increasing. But, for z^n high enough, the consumption bundle $\mathbf{x}(z^n, \tau_t)$ must be close to the autarchy allocation, and by the Inada conditions on utility, this will yield the consumer in the deviating country a lower utility than (for example) the bundle $\mathbf{x}(\tau_t, \tau_t)$ generated by not deviating. Contradiction. \square

Proof of Lemma 2. By definition, $z(\tau) = \max\{\zeta(\tau), \tau\}$. Moreover, as $\zeta(\cdot)$ is decreasing in τ , it must be the case that there exists a $\bar{\tau}$ for which $\zeta(\tau) > \tau$, $\tau < \bar{\tau}$, $\zeta(\tau) < \tau$, $\tau > \bar{\tau}$.

We now prove that $\bar{\tau} < \hat{\tau}$. Suppose not; consider $\bar{\tau} = \hat{\tau}$ first. By the definition of (3.7) we must have $\zeta(\hat{\tau}) = \hat{\tau} = \arg \max_z \{w(\hat{\tau}, \hat{\tau}) + \delta w(\hat{\tau}, \hat{\tau}) / (1 - \delta)\}$. The first order condition requires that

$$w_1(\hat{\tau}, \hat{\tau}) + \frac{\delta}{1 - \delta}(w_1(\hat{\tau}, \hat{\tau}) + w_2(\hat{\tau}, \hat{\tau})) = 0$$

But by a standard argument, the myopic best response tariff $\hat{\tau}$ solves $w_1(\hat{\tau}, \hat{\tau}) = 0$. By A2, we have that $w_1(\hat{\tau}, \hat{\tau}) + w_2(\hat{\tau}, \hat{\tau}) < 0$. Therefore, the first order condition cannot be satisfied at $\bar{\tau} = \hat{\tau}$; a contradiction. Then $\bar{\tau} > \hat{\tau}$ can also be ruled out because $w_1(\bar{\tau}, \bar{\tau}) < 0$ for $\bar{\tau} > \hat{\tau}$.

Combining the fact that $z(\tau) = \max\{\zeta(\tau), \tau\}$ and the fact that there exists a unique $\bar{\tau}$ for which $\bar{\tau} = \zeta(\tau)$, we see that $z(\tau) = \zeta(\tau)$, $\tau < \bar{\tau}$, and $z(\tau) = \tau$, $\tau \geq \bar{\tau}$. \square

Proof of Proposition 2. Suppose to the contrary that $\tau_t = 0$ for some t . Then, at t , the incentive constraint is

$$(1 - \delta)w(0, 0) + \delta w(0, 0) \geq (1 - \delta)w(z(0), 0) + \delta w(z(0), z(0)) \quad (3.22)$$

Now, we will show that at the solution to problem (3.6), $z(0) > 0$. It will then follow that

$$(1 - \delta)w(z(0), 0) + \delta w(z(0), z(0)) > (1 - \delta)w(0, 0) + \delta w(0, 0)$$

contradicting (3.9). To see that $z(0) > 0$, suppose to the contrary that $z(0) = 0$. Note that by the optimality of free trade, $w(0, 0) > w(\tau, \tau)$, $\tau \neq 0$, which of course implies that

$$w_1(0, 0) + w_2(0, 0) = 0$$

Now, consider a small increase in z_t from 0, say Δ . Then, the effect of this change in z_t on the deviation payoff is

$$\Delta [(1 - \delta)w_1(0, 0) + \delta(w_1(0, 0) + w_2(0, 0))] = (1 - \delta)\Delta w_1(0, 0) > 0$$

where the last inequality follows from A1. \square

Proof of Proposition 3. The only part that does not follow directly from Figure 1 is that $\tau^* = \bar{\tau}$. To prove this, it is sufficient to show that on the interval $[0, \bar{\tau}]$, the slope of α is greater than the slope of β in absolute value. This slope condition clearly rules out the case in Figure 1, where $\tau^* < \bar{\tau}$.²¹ Now, the slope of β is

$$\beta'(\tau) = w_1(\tau, \tau) + w_2(\tau, \tau) \tag{3.23}$$

Moreover, from Lemma 2, the constraint $z \geq \tau$ is not binding on $[0, \bar{\tau}]$, so differentiating α and applying the envelope theorem gives:

$$\alpha'(\tau) = (1 - \delta)w_2(z, \tau) \tag{3.24}$$

Given $z \geq \tau$ in (3.24), we must have

$$w_2(z, \tau) - w_2(\tau, \tau) = \int_{\tau}^z [w_{12}] dx,$$

²¹The case shown in Figure 2, where $\tau^* < \bar{\tau}$, requires that the slope of α must be less than that of β in absolute value somewhere in the interval $[\tau^*, \bar{\tau}]$.

and from A3 we have $w_2(z, \tau) - w_2(\tau, \tau) < 0$, so

$$\alpha'(\tau) \leq (1 - \delta)w_2(\tau, \tau). \quad (3.25)$$

So, from (3.23), (3.25), the required condition is that

$$(1 - \delta)w_2(\tau, \tau) < w_1(\tau, \tau) + w_2(\tau, \tau)$$

Rearranging, this is

$$0 < w_1(\tau, \tau) + \delta w_2(\tau, \tau) \quad (3.26)$$

But, the FOC defining τ is:

$$w_1(\bar{\tau}, \bar{\tau}) + \delta w_2(\bar{\tau}, \bar{\tau}) = 0 \quad (3.27)$$

As $\tau < \bar{\tau}$, from (3.27) we must have:

$$w_1(\tau, \tau) + \delta w_2(\tau, \tau) = - \int_{\tau}^{\bar{\tau}} [w_{11} + (1 + \delta)w_{12} + \delta w_{22}] dx \quad (3.28)$$

where the derivatives on the RHS of (3.28) are evaluated at (x, x) . By A3, $w_{12} < 0$.

By assumption, $w_{11}, w_{22} \leq 0$. So, (3.28) implies (3.26), as required.

The fact that $\tau^* = \bar{\tau} < \hat{\tau}$ follows from Lemma 2. \square

Proof of Proposition 4. (a) Following the proof of Lockwood and Thomas (2002), Lemma 2.2, the equilibrium conditions (3.10) can be shown to be equivalent to the following difference equation,

$$\alpha(\tau_{t+1}) = \frac{1}{\delta} [\alpha(\tau_t) - (1 - \delta)\beta(\tau_t)], \quad t = 1, \dots \quad (3.29)$$

with initial condition $\tau_0 = \hat{\tau}$, plus the condition that the solution to (3.29) is bounded. To see this, note first that advancing the equality in (3.29) by one period (i.e. from t to $t + 1$), multiplying the $t + 1$ -condition by δ and subtracting from the t -condition, we get:

$$\begin{aligned} (1 - \delta)w(\tau_t, \tau_t) &= (1 - \delta)w(z(\tau_t), \tau_t) + \delta w(z(\tau_t), (z(\tau_t))) \\ &\quad - \delta [(1 - \delta)w(z(\tau_{t+1}), \tau_{t+1}) + \delta w(z(\tau_{t+1}), (z(\tau_{t+1})))], \quad t = 1, \dots \end{aligned} \quad (3.30)$$

Using the definitions of α, β in (3.30) and rearranging, we get²² (3.29).

(b) Now suppose that the path $\{\tau_t\}$ is in E and more efficient than the stationary path $\bar{\tau}$. Then, for some t , $\tau_t < \bar{\tau}$ (otherwise, $\tau_t \geq \bar{\tau}$, all t , so it cannot be more efficient). We now show that if $\tau_t < \bar{\tau}$, then $\tau_{t+1} < \tau_t$. For suppose not. then, as α is decreasing in τ_t , we would have

$$\alpha(\tau_{t+1}) \leq \alpha(\tau_t) \quad (3.31)$$

Combining (3.29) and (3.31), we have

$$\frac{1}{\delta} [\alpha(\tau_t) - (1 - \delta)\beta(\tau_t)] \leq \beta(\tau_t) \implies \alpha(\tau_t) \leq \beta(\tau_t)$$

But as $\tau_t < \bar{\tau}$, $\alpha(\tau_t) > \beta(\tau_t)$, a contradiction. So, any solution of (3.29) is clearly a strictly decreasing sequence. There are then two possibilities. First, $\lim_{t \rightarrow \infty} \tau_t = \tau_\infty > -\infty$. But then $\alpha(\tau_\infty) = \beta(\tau_\infty)$, contradicting the definition of $\bar{\tau} > \tau_\infty$ as the smallest root of $\alpha(\tau) = \beta(\tau)$. The other is $\lim_{t \rightarrow \infty} \tau_t = -\infty$. But this path cannot be more efficient than the stationary path, a contradiction. \square

Proof of Lemma 4. The proof is by induction. Assume $\tau_t < \tau_{t-1}$. Rewriting (3.15), we get:

$$\begin{aligned} \delta [w(\tau_t, \tau_{t+1}) - w(\tau_t, \tau_t)] &= w(\tau_{t-1}, \tau_t) + \frac{\delta w(\tau_{t-1}, \tau_{t-1})}{1 - \delta} - \delta \left[w(\tau_t, \tau_t) + \frac{\delta w(\tau_t, \tau_t)}{1 - \delta} \right] \\ &= \max_{\tau_t \leq z_t \leq \tau_{t-1}} \left\{ w(z_t, \tau_t) + \frac{\delta w(z_t, z_t)}{1 - \delta} \right\} - \delta \left[w(\tau_t, \tau_t) + \frac{\delta w(\tau_t, \tau_t)}{1 - \delta} \right] \end{aligned}$$

By Lemma 3,

$$\begin{aligned} &w(\tau_{t-1}, \tau_t) + \frac{\delta w(\tau_{t-1}, \tau_{t-1})}{1 - \delta} - \left[w(\tau_t, \tau_t) + \frac{\delta w(\tau_t, \tau_t)}{1 - \delta} \right] \\ &= \max_{\tau_t \leq z_t \leq \tau_{t-1}} \left\{ w(z_t, \tau_t) + \frac{\delta w(z_t, z_t)}{1 - \delta} \right\} - \left[w(\tau_t, \tau_t) + \frac{\delta w(\tau_t, \tau_t)}{1 - \delta} \right] \\ &> 0 \end{aligned}$$

²²The converse result can be obtained by solving (3.29) forward by substitution to get:

$$\alpha(\tau_t) = (1 - \delta)(\beta(\tau_t) + \delta\beta(\tau_{t+1}) + \dots + \delta^n\beta(\tau_{t+n})) + \delta^{n+1}\alpha(\tau_{t+n+1})$$

So, as long as $\lim_{t \rightarrow \infty} \alpha(\tau_t) = 0$, (3.29) implies (3.10).

where the third line follows by definition. And because $0 < \delta < 1$, it follows that $\delta [w(\tau_t, \tau_{t+1}) - w(\tau_t, \tau_t)] > 0$. So, $w(\tau_t, \tau_{t+1}) > w(\tau_t, \tau_t)$. But then, by A1, $\tau_{t+1} < \tau_t$, as required. \square

Proof of Proposition 5. First, rewrite (3.16) as a function of τ_1 :

$$\begin{aligned} f(\tau_1) &= (1 - \delta)w(\chi(\tau_1, \hat{\tau}), \tau_1) + \delta w(\chi(\tau_1, \hat{\tau}), \chi(\tau_1, \hat{\tau})) \\ &\quad - (1 - \delta)(w(\tau_1, \tau_1) + \delta w(\tau_2(\hat{\tau}, \tau_1), \tau_2(\hat{\tau}, \tau_1)) + \dots) \end{aligned}$$

Now, note that by the definition of $\bar{\tau}$,

$$(1 - \delta)w(\chi(\bar{\tau}, \hat{\tau}), \bar{\tau}) + \delta w(\chi(\bar{\tau}, \hat{\tau}), \chi(\bar{\tau}, \hat{\tau})) = w(\bar{\tau}, \bar{\tau})$$

Moreover, $\tau_t(\hat{\tau}, \bar{\tau}) < \bar{\tau}$, all t by Lemma 4. So, if $\tau_1 = \bar{\tau}$, (3.12) is slack i.e.

$$\begin{aligned} (1 - \delta)(w(\bar{\tau}, \bar{\tau}) + \delta w(\tau_2(\hat{\tau}, \bar{\tau}), \tau_2(\hat{\tau}, \bar{\tau})) + \dots) &> w(\bar{\tau}, \bar{\tau}) \\ &= (1 - \delta)w(\chi(\bar{\tau}, \hat{\tau}), \bar{\tau}) + \delta w(\chi(\bar{\tau}, \hat{\tau}), \chi(\bar{\tau}, \hat{\tau})) \end{aligned}$$

where the inequality follows by A2. So, we have shown that $f(\bar{\tau}) < 0$.

Next, if $\tau_1 = \varepsilon$, we have

$$(1 - \delta)w(\chi(\varepsilon, \hat{\tau}), \varepsilon) + \delta w(\chi(\varepsilon, \hat{\tau}), \chi(\varepsilon, \hat{\tau})) = \max_{\varepsilon \leq z \leq \hat{\tau}} (1 - \delta)w(z, \varepsilon) + \delta w(z, z) > w(\varepsilon, \varepsilon)$$

for ε small enough: the inequality is strict by Lemma 2 above, as for ε small enough, $z(\varepsilon) > \varepsilon$. Moreover, from Lemma 4, for ε small enough,

$$(1 - \delta)(w(\varepsilon, \varepsilon) + \delta w(\tau_2(\hat{\tau}, \varepsilon), \tau_2(\hat{\tau}, \varepsilon)) + \dots) \simeq w(\varepsilon, \varepsilon)$$

So, it is possible to choose ε small enough so that

$$(1 - \delta)(w(\varepsilon, \varepsilon) + \delta w(\tau_2(\hat{\tau}, \varepsilon), \tau_2(\hat{\tau}, \varepsilon)) + \dots) < (1 - \delta)w(\chi(\varepsilon, \hat{\tau}), \varepsilon) + \delta w(\chi(\varepsilon, \hat{\tau}), \chi(\varepsilon, \hat{\tau}))$$

i.e. $f(\varepsilon) > 0$. Now, by inspection, $f(\cdot)$ is continuous in τ_1 as χ and τ_t are continuous in τ_1 . So, there exists at least one value of τ_1 for which $f(\tau_1) = 0$, and so there exists a smallest such value. \square

3.6.2 An Example: Quasi-linear Preferences

We assume that the utility function is of quasi-linear form given by (3.2). Maximization of (3.2) subject to (3.3) gives demands for the two goods;

$$x_j^i = \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{-\sigma}, \quad j \neq i \quad (3.32)$$

$$x_i^i = 1 + \frac{R_i}{p_i} - \sum_{j \neq i} \frac{p_j(1 + \tau_j^i)x_j^i}{p_i} = 1 + \frac{R_i}{p_i} - \sum_{j \neq i} \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{1-\sigma} \quad (3.33)$$

where the demand for good i , x_i^i is determined residually via the budget constraint.

Indirect utility for the representative household in i is therefore derived by substituting (3.32), (3.33), back into (3.2) to get

$$v^i = \frac{1}{\sigma - 1} \sum_{j \neq i} \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{1-\sigma} + \frac{R_i}{p_i} \quad (3.34)$$

Also, tariff revenue is

$$R_i = \sum_{j \neq i} p_j \tau_j^i x_j^i = \sum_{j \neq i} \frac{p_j \tau_j^i}{p_i} \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{-\sigma} \quad (3.35)$$

We substitute (3.35) into (3.34) to get:

$$v^i = \frac{1}{\sigma - 1} \sum_{j \neq i} \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{1-\sigma} + \sum_{j \neq i} \frac{p_j \tau_j^i}{p_i} \left[\frac{p_j(1 + \tau_j^i)}{p_i} \right]^{-\sigma} \quad (3.36)$$

Now, in Nash tariff equilibrium, a given country will always set the same tariff on all imported goods. So, we may suppose that all countries $j \neq i$ set a tariff $\tau' = \tau_k^j$ on imports from all countries $k \neq j$, and country i sets tariff $\tau = \tau_k^i$, $k \neq i$. Then, we only need to find the best response τ to τ' to characterize the Nash equilibrium in tariffs. If $\tau' = \tau_{jk}$, $k \neq j, \dots, n$, $\tau = \tau_k^i$, $k \neq i$, then in equilibrium, $p_j = p$, all $j \neq i$. So, we may choose p_i as the numeraire. Using these simplifications, we may rewrite (3.36) as

$$v(\tau, p) = \frac{n-1}{\sigma-1} [p(1+\tau)]^{1-\sigma} + (n-1)p\tau [p(1+\tau)]^{-\sigma} \quad (3.37)$$

Finally, we need to calculate how the (reciprocal of) terms of trade for country i , p , changes with τ', τ . Evaluating (3.32), (3.33) at $\tau' = \tau_{jk}$, $k \neq j, \dots, n$, $\tau = \tau_k^i$, $k \neq i$, $p_j = p$, $j \neq i$, $p_i = 1$, we get;

$$x_i^i = 1 + (n-1)p\tau [p(1+\tau)]^{-\sigma} - (n-1)[p(1+\tau)]^{1-\sigma} \quad (3.38)$$

$$x_i^j = \left[\frac{(1+\tau')}{p} \right]^{-\sigma} \quad (3.39)$$

So, substituting (3.38), (3.39) into the market-clearing condition for good i , namely that supply of unity equals the sum of country demands ($1 = \sum_{i \in N} x_i^j$), we have

$$(n-1)p\tau [p(1+\tau)]^{-\sigma} - (n-1)[p(1+\tau)]^{1-\sigma} + (n-1) \left[\frac{(1+\tau')}{p} \right]^{-\sigma} = 0 \quad (3.40)$$

Solving (3.40) for p , we get:

$$p(\tau, \tau') = \left(\frac{1+\tau}{1+\tau'} \right)^{\sigma/(1-2\sigma)}$$

Note that as $\sigma > 0.5$ by assumption, $p_\tau < 0$ i.e. an increase in i 's tariff always improves i 's terms of trade. So, we may write country i 's indirect utility as

$$w(\tau, \tau') \equiv v(p(\tau, \tau'), \tau) = \frac{n-1}{\sigma-1} [p(1+\tau)]^{1-\sigma} + (n-1)p\tau [p(1+\tau)]^{-\sigma}$$

So, a (symmetric) Nash equilibrium in tariffs is a $\hat{\tau}$ such that $v(\hat{\tau}, p(\hat{\tau}, \hat{\tau})) \geq v(\tau, p(\tau, \hat{\tau}))$, all $\tau \neq \hat{\tau}$.

As v is continuously differentiable, we can characterize $\hat{\tau}$ as the solution to

$$v_\tau(\hat{\tau}, p(\hat{\tau}, \hat{\tau})) + v_p(\hat{\tau}, p(\hat{\tau}, \hat{\tau}))p_\tau(\hat{\tau}, \hat{\tau}) = 0 \quad (3.41)$$

where v_τ, v_p denote partial derivatives of v . Now,

$$v_\tau(\tau, p) = -\sigma(n-1)\tau p^{1-\sigma}(1+\tau)^{-\sigma-1} \quad (3.42)$$

$$v_p(\tau, p) = -(n-1)p^{-\sigma}(1+\tau)^{1-\sigma} + (n-1)(1-\sigma)p^{-\sigma}\tau(1+\tau)^{-\sigma}$$

$$p_\tau = \frac{\sigma}{1-2\sigma} \left(\frac{1+\tau}{1+\tau'} \right)^{(\sigma/(1-2\sigma))-1} \frac{1}{1+\tau'}$$

So, using (3.42) and the fact that $p(\hat{\tau}, \hat{\tau}) = 1$, we have from (3.41) that

$$-\sigma(n-1)\hat{\tau}(1+\hat{\tau})^{-\sigma-1} + [-(n-1)(1+\hat{\tau})^{1-\sigma} + (n-1)(1-\sigma)\hat{\tau}(1+\hat{\tau})^{-\sigma}] \frac{\sigma}{1-2\sigma} \frac{1}{1+\hat{\tau}} = 0$$

Eliminating common terms, we get

$$-\hat{\tau} + [-(1 + \hat{\tau}) + (1 - \sigma)\hat{\tau}] \frac{1}{1 - 2\sigma} = 0$$

Solving, we get

$$\hat{\tau} = \frac{1}{\sigma - 1}$$

for the optimal tariff. Recall that $\sigma > 1$, so $\hat{\tau}$ is defined and positive.

Now we have $\hat{\tau}$, we can check that A1, A2 and A3 hold for tariffs set on the interval $[0, \hat{\tau}]$

Substituting for $p(\tau, \tau')$, we can write the payoff function as follows:

$$w(\tau, \tau') = (n - 1) \left(\frac{(1 + \tau)^{1-\sigma}}{\sigma - 1} + \tau(1 + \tau)^{-\sigma} \right) \left(\frac{1 + \tau}{1 + \tau'} \right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

We can use this expression to verify that A1, A2 and A3 hold. Take A1 first:

$$w_1(\tau, \tau') = (n - 1) \frac{\sigma(1 + \tau)^{-1-\sigma}(1 - (\sigma - 1)\tau)}{2\sigma - 1} \left(\frac{1 + \tau}{1 + \tau'} \right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

The sign of this expression depends on the term in brackets $(1 - (\sigma - 1)\tau)$. If $\tau = \hat{\tau} = 1/(\sigma - 1)$ and $(1 - (\sigma - 1)\tau) = 0$ so $w_1(\tau, \tau') = 0$. If $\tau < \hat{\tau}$ then $(1 - (\sigma - 1)\tau) > 0$ and so $w_1(\tau, \tau') > 0$ as required.

$$w_2(\tau, \tau') = -(n - 1) \frac{\sigma(1 + \tau)^{-1-\sigma}(1 + \sigma\tau)}{2\sigma - 1} \left(\frac{1 + \tau}{1 + \tau'} \right)^{(1-\sigma-\sigma^2)/(1-2\sigma)} < 0 \text{ for all } \tau, \tau' \geq 0.$$

Now A2:

$$w_1(\tau, \tau') + w_2(\tau, \tau') = -(n - 1) \frac{\sigma(1 + \tau)^{-2-\sigma}(\sigma\tau(2 + \tau + \tau') - (1 + \tau)\tau')}{2\sigma - 1} \left(\frac{1 + \tau}{1 + \tau'} \right)^{(1-\sigma-\sigma^2)/(1-2\sigma)}$$

Now the sign of this expression depends on the term in brackets $(\sigma\tau(2 + \tau + \tau') - (1 + \tau)\tau')$.

It is easy to see that when $\tau = \tau' = 0$ we have $(\sigma\tau(2 + \tau + \tau') - (1 + \tau)\tau') = 0$ and therefore $w_1(\tau, \tau') + w_2(\tau, \tau') = 0$. This is necessary for free trade to maximize efficiency. Moreover, by inspection $(\sigma\tau(2 + \tau + \tau') - (1 + \tau)\tau') > 0$ for all

$\tau, \tau' \in (0, \hat{\tau})$, $\sigma > 1$, so $w_1(\tau, \tau') + w_2(\tau, \tau') < 0$ as required. Finally, regarding A3:

$$w_{12}(\tau, \tau') = -(n-1) \frac{(\sigma-1)\sigma^2(1+\tau)^{-2-\sigma}(1-(\sigma-1)\tau)}{(2\sigma-1)^2} \left(\frac{1+\tau}{1+\tau'}\right)^{\sigma(1-\sigma)/(1-2\sigma)}.$$

So $w_{12}(\tau, \tau') < 0$ because $(1-(\sigma-1)\tau) > 0$ for $\tau, \tau' \in (0, \hat{\tau})$ as required.

Now we want to characterize the constrained deviation, using it to derive $\bar{\tau}$. Dropping time subscripts and setting this first order condition equal to zero, we have

$$w_1(z(\tau), \tau) + \frac{\delta}{1-\delta}(w_1(z(\tau), z(\tau)) + w_2(z(\tau), z(\tau))) = 0.$$

We can write (3.2) as follows

$$w(z(\tau), \tau) = (n-1) \left(\frac{1+z(\tau)}{1+\tau}\right)^{\sigma(1-\sigma)/(1-2\sigma)} \beta(z(\tau)),$$

where $\gamma(z(\tau)) = \frac{(1+z(\tau))^{1-\sigma}}{\sigma-1} + z(\tau)(1+z(\tau))^{-\sigma}$, so $\gamma'(z(\tau)) = -\sigma z(\tau)(1+z(\tau))^{-1-\sigma}$. Then

$$w_1(z(\tau), \tau) = \frac{\frac{\sigma(1-\sigma)}{1-2\sigma} w(z(\tau), \tau)}{(1+z(\tau))} + (n-1) \left(\frac{1+z(\tau)}{1+\tau}\right)^{\frac{\sigma(1-\sigma)}{1-2\sigma}} \gamma'(z(\tau)),$$

and

$$w_2(z(\tau), \tau) = -\frac{\frac{\sigma(1-\sigma)}{1-2\sigma} w(z(\tau), \tau)}{(1+\tau)}$$

It is then straightforward to see that the first order condition can be rewritten $(1-\delta)w_1(z(\tau), \tau) + \delta\gamma'(z(\tau)) = 0$. Setting $z(\tau) = \tau = \bar{\tau}$ in the first order condition, we get

$$(1-\delta) \frac{\sigma(\sigma-1)}{2\sigma-1} \frac{\gamma(\bar{\tau})}{1+\tau^*} + \gamma'(\bar{\tau}) = 0$$

Substituting for $\gamma(\bar{\tau})$ and $\gamma'(\bar{\tau})$ and simplifying, the equation becomes

$$\frac{\sigma(1+\bar{\tau})^{-1-\sigma}(1-\delta+(1-\sigma(1+\delta))\bar{\tau})}{2\sigma-1} = 0$$

Solving, the only admissible root²³ is

$$\bar{\tau} = \frac{1-\delta}{\sigma(1+\delta)-1}.$$

²³The root $\tau = -1$ also solves this expression.

Figure 1

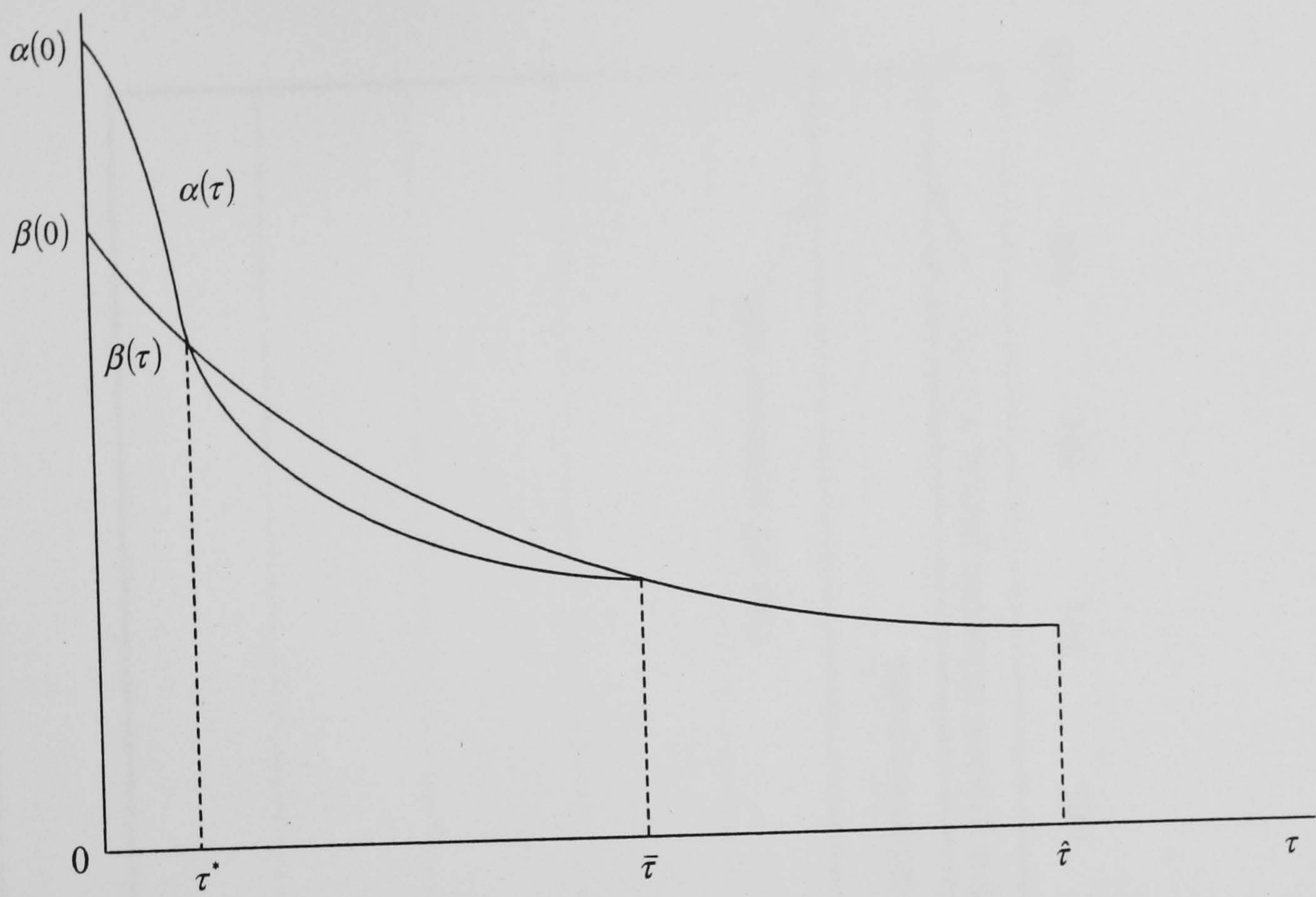


Figure 2

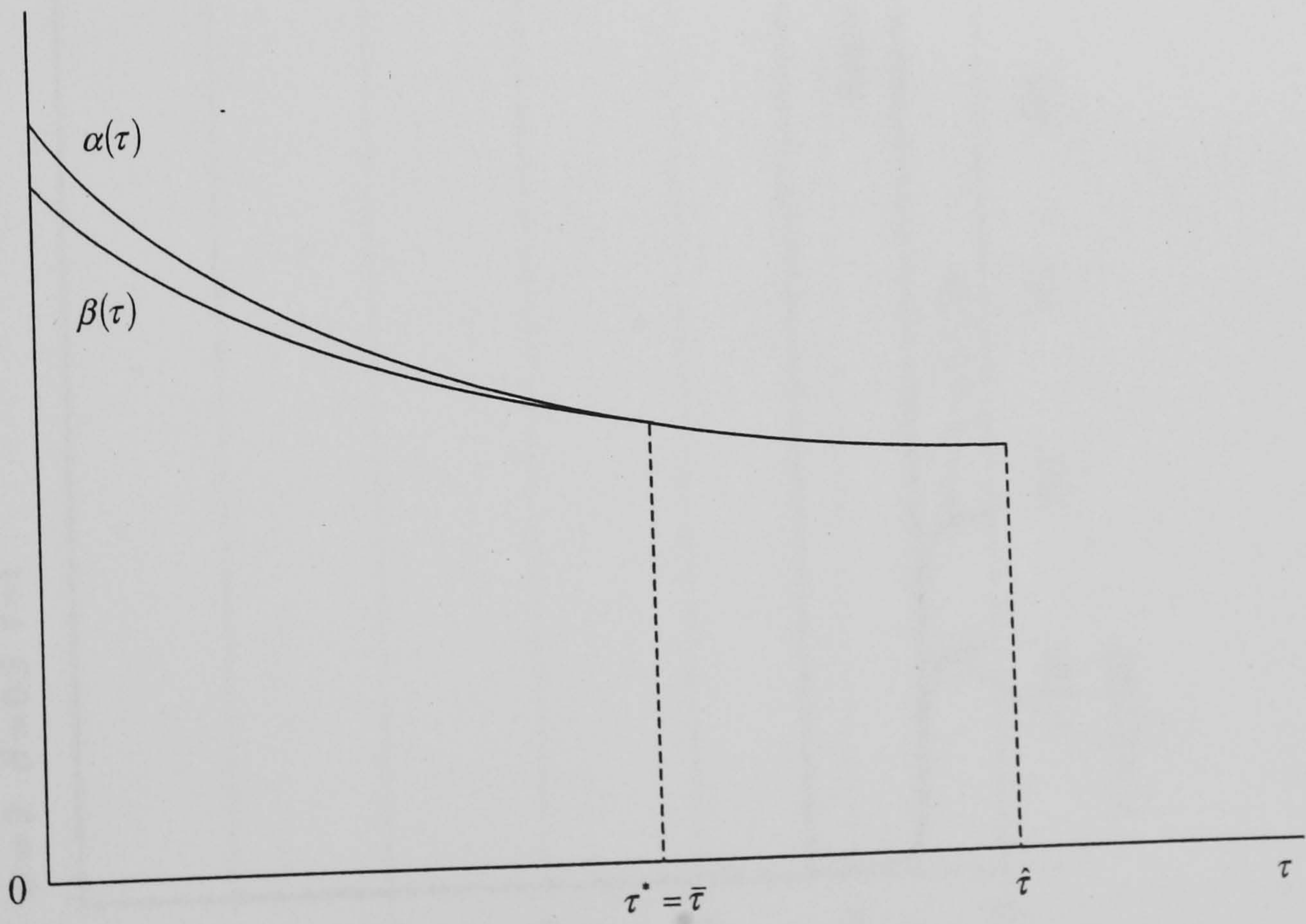


Figure 3; Approximating the Optimal Tariff Reduction Path

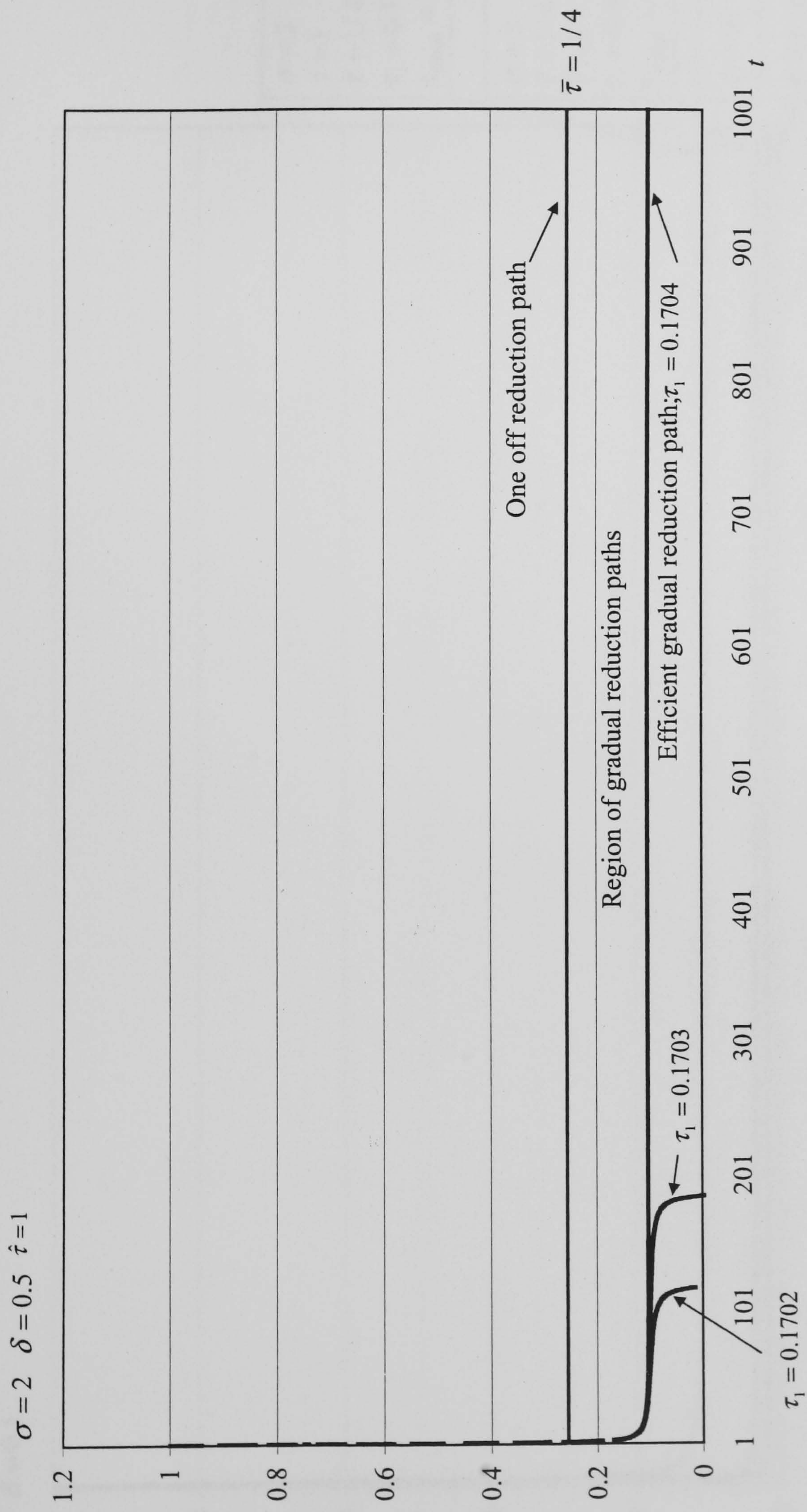
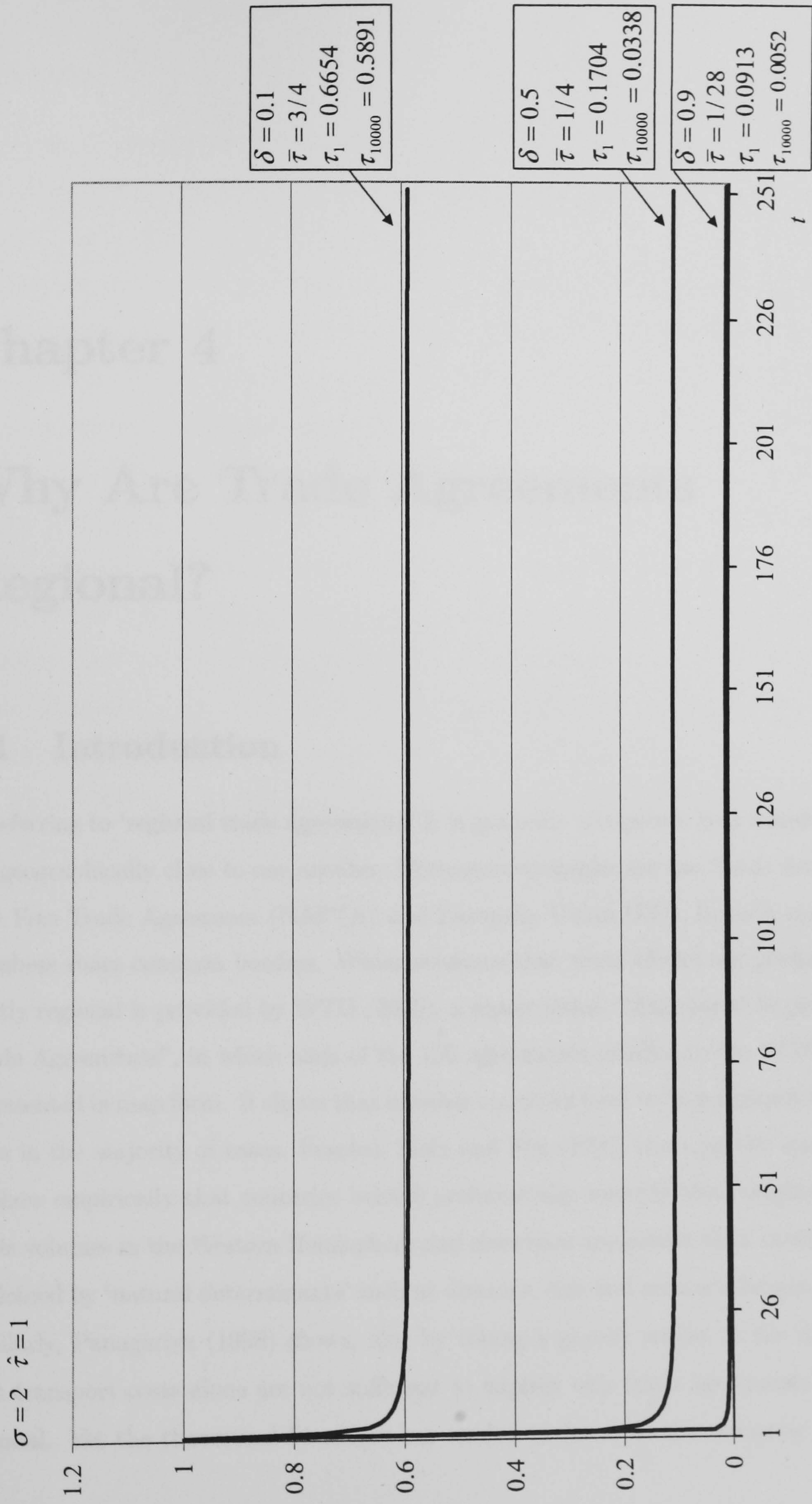


Figure 4: The approximate optimal tariff reduction path for various substitution elasticities



Figure 5: The approximate optimal tariff reduction path for various discount rate:



Chapter 4

Why Are Trade Agreements Regional?

4.1 Introduction

In referring to ‘regional trade agreements’ it is generally recognized that members are geographically close to one another. Prominent examples are the North American Free Trade Agreement (NAFTA) and European Union (EU). In both cases, members share common borders. Wider evidence that trade blocks are predominantly regional is provided by WTO (2000), a report titled “Mapping of Regional Trade Agreements”, in which each of the 150 agreements notified to the WTO is represented in map form. It shows that member countries tend to be geographically close in the majority of cases. Frankel, Stein and Wei (1995) use a gravity model to show empirically that countries behave preferentially towards close neighbors; trade volumes in the Western Hemisphere and elsewhere are greater than could be explained by ‘natural determinants’ such as distance, size and common languages. Similarly, Panagariya (1998) shows, also by taking a gravity model to the data, that transport costs alone are not sufficient to explain why trade agreements are regional. Yet the theoretical literature has tended to focus on the economic im-

plications of regional trade agreements. Almost the entire literature leaves aside the question of *why* there may be strategic incentives to form regional trade agreements.

The purpose of this paper is to argue that politicians balance the increased likelihood that they will be voted for if they coordinate an FTA against the coordination cost itself. The increased likelihood of being voted for results from conventional welfare gains to trade due to formation of the FTA. In the model, trade based gains to an agreement with (close) countries of the same region are higher than gains to an agreement involving (distant) countries from different regions. The costs of bringing politicians together from different nations in order to coordinate an agreement are assumed to be proportional to the number of countries involved, and not dependent upon which countries the politicians come from. Therefore, a regional FTA may be worth coordinating whilst one involving countries from outside the region may not.¹

What is the basis for higher production-trade payoffs to a regional agreement? According to standard optimal tariff theory, the higher the rents made by a foreign

¹Other papers in the literature have had similar concerns to the present paper, or used modes of analysis that are technically similar. Bond (1999) is closest in the question that he addresses. He compares the sustainability of multilateral versus regional trade agreements in a repeated game setting, where both types of agreement are sustained through trigger strategies. In Bond's model, optimal tariffs are higher between closer neighbors, and this makes regional agreements easier to sustain using trigger strategies. Whilst some of Bond's results are related, his approach is quite different, not using profit shifting to motivate tariffs, nor the notion of non-cooperative networks to determine equilibrium. The approach of the present paper allows a wider range of dynamic equilibria to be characterized, as discussed below. Other papers, by Goyal and Joshi (2000) and Furusawa and Konishi (2002) are technically similar, in modelling trade agreements as networks. Both papers show that free trade will not necessarily arise. In the case of Goyal and Joshi (2000) this is due to coordination failure. Furusawa and Konishi (2002) show that free trade fails when countries form customs unions. Both papers take a cooperative rather than a non-cooperative approach to the modelling of trade agreements as networks, and neither paper has a regional dimension.

firm in the domestic market, the more scope there is for shifting rents to domestic citizens through the use of higher tariffs. And because trading costs increase with distance, firms make higher rents in nearby markets than those that are further away. So in the absence of an agreement, optimal tariffs are higher on imports from countries in the same region than on imports from countries of other regions. It follows that a bilateral free trade agreement (FTA) between two close neighbors brings about larger production and trade gains than between distant countries because the former entails a larger mutual tariff reduction.

Whilst standard optimal tariff theory provides a basis for individual tariff setting, a general framework is needed in which the overall structure of trade agreements in the (world) economy can be analyzed. Politicians' incentives to form international trade agreements throughout the world is formalized by adapting Bala and Goyal's (2000) model of noncooperative network formation. Bala and Goyal bring the communication networks previously modelled by others, notably Myerson (1977) and Jackson and Wolinski (1996), into a noncooperative setting.

In communications networks, players benefit from being linked to each other directly and indirectly. For example, if you know someone is the friend of a friend, you can ring up your mutual friend for their phone number. As pointed out by other researchers previously, when communications networks are formed on a cooperative basis they can suffer from coordination failures. The problem is illustrated most clearly in the present setting of trade agreements by Goyal and Joshi (2000). They model FTAs in the manner of a communication network where network formation is cooperative. As a result, whilst free trade is the most efficient Nash equilibrium, it is by no means unique. Other less efficient FTA structures can be an equilibrium because countries may simply fail to coordinate on membership. In the present context it is important to rule out such possibilities. Otherwise it would be possible to have equilibria with only regional trade blocks resulting from nothing more than failures of coordination.

Bala and Goyal address the problem of coordination failure by making indi-

vidual agents responsible for the cost of coordinating a network. Through their sponsorship, individual agents can form a network if it is in their interest without being encumbered by the need to coordinate with other agents. Then a Nash network is one where no agent can do any better by sponsoring any other network or withdrawing their support for the networks that they sponsor, taking as given networks that they do not sponsor.

By taking a noncooperative network approach, coordination failures are ruled out as a possible cause for regional FTAs. We will say that in each period, each FTA must have a sponsor. A sponsor is the country that meets the cost of bringing all other country representatives to the negotiating table in order to make the agreement for that period. It will be assumed that an FTA cannot be made binding indefinitely. It may be that a government can only credibly commit to an agreement for the duration of its parliament. An example of where a country plays such a leadership role in coordinating such agreements is the country that holds the Presidency of the European Parliament.² Conditions are derived under which, in each period, there is an incentive for some country to step forward as sponsor. Moreover, if a country undertakes to sponsor an FTA, all the proposed partners accept because they anticipate (and realize) production-trade gains.

The main result of the paper concerns the characterization of the equilibrium FTA structure that emerges over time under different levels of sponsorship cost. Not surprisingly, if sponsorship costs are above a certain level then no FTAs will

²The Presidency periodically rotates around country members, fulfilling the role of sponsor in the European Union. Whilst the sequence of presidencies of the European Parliament is agreed before hand, the choice of sponsor is not modelled in this paper. In real life, we often observe a number of countries contributing to the sponsorship costs of an agreement. However, in such situations an agreement may be vulnerable to breakdown because countries free ride on each others' willingness to provide sponsorship. The present paper concentrates on the clear cut case in which a single sponsor steps forward in equilibrium. Alternatively, imagine that all countries do contribute to the agreement but that a leader contributes slightly more. Then contributions by all other countries would be normalized to zero.

form at any point on the equilibrium path, and if they are below a certain level then world free trade will emerge straight away. It is when sponsorship costs are at an intermediate level that regionalism arises and can persist over time. Perhaps most interesting of all, a range of sponsorship costs is identified at which regionalism emerges first before free trade can be reached, providing an answer to Bhagwati's (1992) famous question, "Are trade blocks stepping blocks are stumbling blocks in the path to free trade?"

It is worth emphasizing that the facility to analyze gains from network formation across different types of player, coupled with the fact that differing gains are derived from the structure of an underlying micro-model, is a new development of the present paper. The greater benefits to a regional agreement are *not* simply assumed. They are derived from the mutual removal of relatively high tariffs. The paper develops a way of linking these different micro-founded gains to the payoff structure of a network formation game.

This approach to the analysis of network formation with different types of player potentially makes it possible to study a range of different situations that are of interest in economics. Perhaps the best known example is due to Coase (1960), who points out that firms exerting relatively large externalities on one another are better candidates for mergers motivated by internalization. This situation examined by Coase mirrors that analyzed in the present paper in that different types of player exert externalities of differing size on one another. The substantive difference is that the externality discussed by Coase is environmental rather than terms-of-trade based.

Other papers have examined political incentives to form FTAs. Richardson (1993) considers the effect of an FTA in a setting where governments maximize a political support function which gives added weight to export- and import-competing producers' interests. Grossman and Helpman (1995b) use a common agency framework to analyze the way that contributions from interest groups to the government shapes the formation of FTAs. Both models have similarities to

the one presented here in that gains to trade are similarly assumed to affect the incumbent government's likelihood of being voted for. But Grossman and Helpman show that the government will be pulled away from the outcome that voters prefer by financial contributions from interest groups. Levy (1997) studies politicians' incentives to pursue preferential as against multilateral trade agreements in a median voter framework. Voters are affected differently by various different trade agreements depending on their endowments of capital. The government picks the type of trade agreement, preferential or multilateral, that secures victory in the next election. None of these previous papers considers the incentives to form regional trade agreement.

The paper proceeds as follows. In the next section transport costs are introduced to a model of production and trade. This is then used to derive optimal tariffs which vary according to the distance between countries. Section 3 sets up the model of regions and trade agreements as a noncooperative network, allowing the payoffs of network formation to vary depending on the distance between members. Section 4 then establishes the main results of the paper for a simplified three region model. It is here that the possibility of regional trade agreements is demonstrated, as well as the fact that trade blocks can be stepping blocks to free trade. Section 5 concludes.

4.2 A Model of Optimal Tariffs where Distance Matters

The purpose of this section is to present a model of tariff setting which exhibits the property that distance between countries has an effect on the optimal level of protectionism. In particular, it will be shown that optimal tariffs are higher between close neighbors. As in Brander and Spencer (1984) tariffs shift profits from the foreign firm to the domestic consumer. With lower transport costs between

close neighbors, more rents can be shifted through the use of tariffs. So unlike in conventional models, where each country sets a common tariff on all others, in the present model (in equilibrium) each country sets tariffs that vary, and are declining with distance.

4.2.1 Country Location in Regions

The set $\mathcal{N} = \{1, \dots, n\}$ of countries is finite, with the number of countries being given by n . The *regional structure* $\mathcal{R} = \{R_1, R_2, \dots, R_m\}$ partitions \mathcal{N} into regions, where a *region* is a set $R_k \subseteq \mathcal{N}$: $R_i \cap R_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^m R_i = \mathcal{N}$. Each region is assumed to have the same number of countries in it; $|R_i| = r$, for all $R_i \in \mathcal{R}$, and $r > 1$.³ To avoid trivialities, there is more than one region; $|\mathcal{R}| > 1$.

To make the differences between intra-regional versus inter-regional trade concrete, suppose that each Country $i \in \mathcal{N}$ can be located by the coordinates (x_i, y_i) .⁴ Therefore, the *distance* d_{ij} between any two countries i and j can then be measured by a (Euclidean) distance function.

In order to make precise the distinction between countries by region, assume $(x_i, y_i) = (x_j, y_j)$ for $i, j \in R_k$, $i \neq j$; all countries in the same region have the same location. Also assume that $(x_i, y_i) \neq (x_j, y_j)$ for all $i \in R_i$, $j \in R_j$, $i \neq j$. Assume that $d_{ij} = d_{ji} \geq d > 0$ for all $i \in R_i$, $j \in R_j$, $i \neq j$, and that d_{ij} is finite ($d_{ij} = d_{ji} = 0$ for $i, j \in R_k$).

If the distance relationship between countries across regions has some regularity to it, being based on a regular shape for example, then it will help to be able to summarize the information on distances between countries. Let δ_{ki} be the number of countries at a distance d_k from i . Then, for Country $i \in R_i$, let $D_i = \{(d_1, \delta_{1i}), \dots, (d_z, \delta_{zi})\}$ be the set of pairs (d_k, δ_{ki}) , where there exists at least one

³If regions were of different size, then it might be possible for a country might prefer to form a block with a FTA in another region than with countries in its own region. More will be said about this later.

⁴That is, each country can be located in Euclidean \mathcal{R}_2 -space.

(other) Country $j \in R_j$, $i \neq j$, for which $d_{ij} = d_k$.

Example: Three Regions on an Equilateral Triangle

To keep the analysis relatively simple, the main results of the paper will be established using a three region model. Three regions are enough to capture the interactions we are interested in whilst avoiding extensive notation. A three region model is obviously appealing because it captures the interactions between the three most important regions in economic terms, The Americas, Europe and Asia.

To fix ideas, consider the three region example ($|\mathcal{R}| = 3$), with $r = 3$, where each region is located at a distinct vertex of an equilateral triangle. Using an equilateral triangle simplifies the analysis because (all else equal) the payoff to a country of forming an agreement with countries outside its own region is the same, as they must all be at the same distance, regardless of which other region(s) they are in. Label the regions R_a , R_b and R_c . Consider Country i located in region R_a . If the sides of the triangle are of length d , then countries in R_b and R_c are all at distance d from Country i . Then the set D_i has a single element, $D_i = \{(d, 6)\}$, where $(|\mathcal{R}| - 1)r = 6$ gives the number of countries not in R_a .

4.2.2 Production and International Trade with Distance

The specification of demands and production is adapted from Grossman and Helpman (1994, 1995a, 1995b). Each country has a population of size 1. Individuals across countries have identical preferences:

$$u_i(x_{i0}, x_i) = x_{i0} + e \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2, \quad (4.1)$$

where x_{i0} is consumption of the numeraire and x_{ij} denotes consumption in Country i of a good produced in Country j . Citizens across countries are endowed with sufficient quantities of the numeraire to ensure that international markets clear in

equilibrium. The preferences give rise to inverse demands in Country i ,

$$p_{ij} = \frac{du}{dx_{ij}} = e - x_{ij}, \quad (4.2)$$

for goods $j \in \mathcal{N}$, where p_{ij} is the price of Good j in Country i . This form of preferences is clearly restrictive, as there is no substitutability between goods. When combined with the production technology introduced below, the resulting model has the advantage that trade block formation is always trade creating; it yields positive gains through production and trade. As we shall see, these gains can then be balanced against the (political) costs of forming an agreement. More general functional forms would also allow for the possibility of trade diversion. But such situations are not of direct interest here as countries would not consider formation of a trade agreement when trade diversion dominates.⁵

The representative citizen in each Country j is uniquely endowed with a specific factor j that enables him to produce Good j . When referring to him in his productive role, we will refer to the citizen as ‘producer’. The cost of producing a unit of j for sale in Country i is given by the function

$$c_{ij} = c + t_{ij} + d_{ij}, \quad (4.3)$$

where c is the basic per-unit production cost, which is the same for all firms, and t_{ij} is the tariff levied by Country i on imports from Country j . Producers take prices and all elements of costs as given and choose outputs competitively to maximize profits across markets. Also, domestic firms compete perfectly to transport goods to the home market, so that they deliver at cost.⁶

⁵Moreover, as for the contexts examined by Grossman and Helpman, more general functional forms make for an unworkable analytical framework. In a companion paper, Zissimos (2003) I allow Cournot competition in homogeneous products and find that welfare analysis of FTA formation is intractable. Elsewhere in the literature on FTAs, the utility function of Ottaviano, Tabuchi and Thisse (2002) has been adopted. However, in the present context this utility function raises similar problems of tractability.

⁶In a symmetric model, an equivalent assumption would be that the world market for transportation is competitive and that each firm from every country has an equal share of the market.

Since the domestic price of Good 0 is normalized to 1, j earns profit

$$\pi_{ij} = (p_{ij} - c_{ij}) x_{ij} \quad (4.4)$$

on output x_{ij} for Country i . Using (4.2) in (4.4), and expanding,

$$\pi_{ij} = ex_{ij} - x_{ij}^2 - c_{ij}x_{ij}.$$

The function π_{ij} is thus differentiable and strictly concave because $-x_{ij}^2$ is concave, and so j 's problem has a unique maximum.

Producer j 's first order condition in Country i is thus given by

$$\frac{\partial \pi_{ij}}{\partial x_{ij}} = e - 2x_{ij} - c_{ij} = 0,$$

or equivalently,

$$p_{ij} - c_{ij} - x_{ij} = 0. \quad (4.5)$$

We can rearrange (4.5) to get

$$x_{ij} = p_{ij} - c_{ij}$$

From this, note the following convenient property; in equilibrium, profits can be written as

$$\pi_{ij} = (p_{ij} - c_{ij}) x_{ij} = x_{ij}^2. \quad (4.6)$$

The fact that profits can be represented in this way arises as a result of the linear structure of the model.

Use (4.3) and (4.2) in (4.5), then rearrange to get the solution for x_{ij} :

$$x_{ij} = \frac{e - c - t_{ij} - d_{ij}}{2}, \text{ all } i, j \in \mathcal{N} \quad (4.7)$$

As citizen j is uniquely endowed with the specific factor required to produce Good j , the monopoly solution solves his problem in each market i . The novel aspect

of this solution is the fact that it provides a way to take general account of the transportation cost from Country i to Country j .⁷

Note from (4.7) that x_{ij} is decreasing in t_{ij} and d_{ij} . To maintain the assumption that all $j \in \mathcal{N}$ are active on all markets, $e - c$ can be made large enough to ensure that $x_{ij} > 0$. For the domestic market, $t_{ij} = d_{ij} = 0$. Therefore, the weakest possible condition necessary and sufficient to ensure strictly positive output by the domestic firm for the domestic market is $e - c > 0$. This condition will be assumed to hold throughout.

4.2.3 Production-Trade Payoffs

The payoffs to the FTA formation game depend directly on the structure of trading arrangements, that is tariff setting across all countries, and the reciprocal impact on production. For this reason, gains to production and trade will be referred to as production-trade payoffs. They are given this name to distinguish them from (net) payoffs to FTA formation once the cost of sponsoring agreements is taken into account.

The representative citizen in Country i receives his production-trade payoff through five economic components: domestic consumer surplus (CS_i), the domestic firm's profit at home and abroad (π_{ii} and π_{ji} , $j \neq i$ respectively), tariff revenue (TR_i), and net profits from transportation (DR_i):

$$w_i = CS_i + \pi_{ii} + \sum_{j \in \mathcal{N}/\{i\}} \pi_{ji} + TR_i + DR_i. \quad (4.8)$$

The optimal tariff \hat{t}_{ij} is derived by maximizing this expression with respect to t_{ij} . To do this, w must be expressed in terms of model variables; the subject of the next result.

⁷Two country models where production takes account of iceberg transportation costs are commonplace in the literature. The novel aspect of this solution is that it allows for quantities produced for any number of trade partners at any number of distances to be calculated explicitly.

Lemma 1. Let $CS_i = \sum_{j \in \mathcal{N}} \frac{1}{2} (e - p_{ij}) x_{ij}$, $TR_i = \sum_{j \in \mathcal{N}/\{i\}} t_{ij} x_{ij}$, and $DR_i = \sum_{j \in \mathcal{N}/\{i\}} d_{ij} x_{ij}$. Then

$$\begin{aligned} w_i &= CS_i + \pi_{ii} + \sum_{j \in \mathcal{N}/\{i\}} \pi_{ji} + TR_i + DR_i. \\ &= (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 - \sum_{j \in \mathcal{N}/\{i\}} x_{ij}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2. \end{aligned}$$

With payoffs of the representative citizen in Country i as given by Lemma 1, it is straightforward to solve for the *optimal tariff*; the tariff that maximizes the representative citizen's payoff in a one-shot game with no communication.

4.2.4 Optimal Tariffs with Distance

The solution to the optimal tariff problem is given in the following result:

Proposition 1. The unique optimal tariff set by Country i on imports from Country j takes the form

$$\hat{t}_{ij} = \frac{e - c}{3} - d_{ij}.$$

The key thing to notice is that the optimal tariff is decreasing in distance. The closer a country is, the higher the optimal tariff levied on its imports. The intuition is simple. Higher rents are made in nearby markets because a smaller share of revenue is lost in transportation costs to serve those markets. Consequently, there are more rents available to shift to domestic consumers using the tariff.⁸

As was the case in the solution for x_{ij} given by (4.7), it is always possible to set $e - c$ high enough to ensure that $\hat{t}_{ij} > 0$. If $\hat{t}_{ij} < 0$ then the optimal trade intervention is a subsidy. In that case, free trade is not necessarily welfare maximizing. In the present analysis we will be focusing on the standard case where free trade is best. For the optimal tariff to be positive, it is necessary and sufficient to make the following assumption:

⁸Note that the term $(e - c)/3$ is parametric, and depends on the structure of linear-quadratic demands.

A1. $0 < d_k \leq (e - c) / 3$ for all $d_k \in D_i$, all $i \in \mathcal{N}$.

Assumption A1 holds throughout.

4.2.5 Free Trade Agreements (FTAs)

Having defined optimal tariffs, we are now ready to specify the formal assumptions that define an FTA. If countries i and j agree to adopt free trade - setting $t_{ij} = t_{ji} = 0$ - then they are said to have a *Free Trade Agreement* (FTA). In the absence of an FTA between countries i and j , let Country i set optimal tariffs on imports from Country j and vice versa - $t_{ij} = \hat{t}_{ij} = (e - c) / 3 - d_{ij}$; $\hat{t}_{ji} = \hat{t}_{ij}$.⁹

It is well known from the literature on trade agreements that the formation of an FTA may increase or reduce welfare depending on whether trade creation or trade diversion dominates. The following result shows that in the present model the production-trade gains to FTA formation are always welfare improving for any partner; trade creation dominates.

Proposition 2. Assume A1.

(i) The production-trade gain to two countries i and j from a bilateral FTA is given by

$$\Delta w_i = \frac{1}{72} (7(e - c) + 3d_{ij}) ((e - c) - 3d_{ij})$$

and is positive.

(ii) production-trade gains from a bilateral FTA are decreasing in the distance between members i and j : $d(\Delta w_i) / d(d_{ij}) = -\frac{1}{4}(e - c + d_{ij})$.

Because Country j is the world's only exporter of Good j , there can be no trade diversion associated with trade in Good j when Country j forms a trade agreement with any other country.¹⁰ Proposition 2(i) tells us that there are production-trade

⁹Note that $\hat{t}_{ji} = \hat{t}_{ij}$ because $d_{ji} = d_{ij}$.

¹⁰Recall that trade diversion arises when the formation of an FTA between Countries i and j reduces efficiency because the agreement causes i to import more of Good j from Country j

gains to a trade agreement between two countries. Proposition 2(ii) says that production-trade gains are higher the closer the two countries are to one another.

To see why Proposition 2(ii) holds recall that, in the absence of a trade agreement, Proposition 1 shows that tariffs between regional members are higher than between countries of different regions. Because the distance between two countries i and j in the same region is equal to zero, by Proposition 1, optimal tariffs on mutual trade are $\hat{t}_{ij} = \hat{t}_{ji} = (e - c) / 3$. On the other hand, because there is a positive distance $d > 0$ between two countries i and k in different regions, by Proposition 1, optimal tariffs on mutual trade are $\hat{t}_{ik} = \hat{t}_{ki} = (e - c) / 3 - d$. Proposition 2(ii) shows that, due to the removal of higher tariffs, FTA formation between regional members yields higher production-trade gains than FTA formation between countries of different regions.

Proposition 2 focuses on production-trade gains to an FTA between two countries. As we shall see in Lemma 4, the production-trade payoff to forming a multilateral FTA is a multiple of the production-trade payoff to a bilateral agreement. Due to the fact that there is no trade diversion, and that the firm in each country has a monopoly over the good that it exports, the production-trade gains to trade reaped by each citizen-producer can be calculated as a fixed multiple of the number of countries in the agreement.

If production-trade gains were all that mattered, then Proposition 2 suggests the world would move straight to free trade. Anecdotal discussions often reflect surprise that the process of regionalism has not led more quickly towards free trade. One explanation is that the costs of coordinating such agreements hold the process back. Sponsorship costs, formalized in the next section, play exactly this role in the present model. But by themselves such costs do not explain why FTAs are formed within a region. Proposition 2 indicates that greater benefits to regional (than non-regional) FTA formation can be set against the costs of sponsoring an

when Country $k \neq j$ is a more efficient producer of Good j . In the present model, trade diversion is ruled out by the assumption that Country j is the sole producer of Good j .

agreement, and may make regional FTA formation worthwhile even though non-regional FTA formation is not. This is the central insight that will be developed in the following sections.

4.3 Regions and Trade Agreements in a Non-cooperative Network

This section formalizes FTA formation in a world of regions. It shows how the formal language of noncooperative network formation due to Bala and Goyal (2000) can be adapted to model FTA formation when there is a regional dimension to the model. The distinction that countries make between FTA members in their own region and those from other regions is formalized by an adaptation of Slikker and van den Nouweland's (2000) partitioning of players in a network formation game.

Gains from trade through the formation of an FTA provide a concrete source of benefits from network formation in the model. If countries are not members of the same FTA, they set tariffs non-cooperatively, producing and trading a smaller amount than if they engaged in free trade. If countries form an FTA (network) then they remove tariffs, giving rise to production trade gains.

4.3.1 FTAs as Networks

The overall FTA structure is described by the *graph* (\mathcal{N}, g) , a pair of disjoint sets, where g is a set of links called a (directed) *network*. The FTA structure of each region is described by the *subgraph* (R_k, g_{R_k}) , where g_{R_k} is a set of links called a *subnetwork* between countries of region R_k .

A *strategy* of Country $i \in \mathcal{N}$ is a row vector $g_i = (g_{i1}, \dots, g_{in})$ where $g_{ij} \in \{0, 1\}$. If Country i sponsors an FTA with a set of other countries $\mathcal{A}_j = \{j_1, \dots, j_l\}$ then it sets $g_{ij_k} = 1$ for all $j_k \in \mathcal{A}_j$. For all countries j with which Country i does not sponsor an agreement, it sets $g_{ij} = 0$. If $g_{ij_k} = 1$ (or if $g_{j_k i} = 1$) then there is said

to be a *link* between i and j_k . The strategies of all countries forms $g = \{g_1, \dots, g_n\}$. Since \mathcal{N} can be partitioned into regions, we can also have $g = \{g_{R_1}, \dots, g_{R_m}\}$, where g_{R_k} is the set of links between members of region R_k .

The set of all strategies of Country i is denoted by \mathcal{G}_i . Attention is restricted to pure strategies. The set $\mathcal{G} = \mathcal{G}_1 \times \dots \times \mathcal{G}_n$ is the space of pure strategies of all the countries. This is needed for specification of the payoff function below.

There is a *path* from i to j_k in g if i and j are linked, or if there exist countries i_1, \dots, i_m distinct from each other such that they are all linked. A path in g between i and j is denoted $i \xleftrightarrow{g} j$.¹¹

A set $\mathcal{A}_k \subseteq \mathcal{N}$ is a *component* of g if for all i and j in \mathcal{A}_k there is a path between them, and there does not exist a path between a country in \mathcal{A}_k and one in $\mathcal{N} \setminus \mathcal{A}_k$.

A network g is called *connected* if it has a unique component \mathcal{A} , with all $i \in \mathcal{A} = \mathcal{N}$. A network that is not connected is referred to as *disconnected*. A network is called *empty* if $g_{ij} = g_{ji} = 0$ for all $i, j \in \mathcal{N}$.

The following definitions identify the network structures just introduced to the various types of FTA that will be of interest to us.

Definition 1. (FTA Membership) A component $\mathcal{A}_k \subseteq \mathcal{N}$ of g is an FTA: for all $i \in \mathcal{A}_k$, if $j \in \mathcal{A}_k$ then $t_{ij} = 0$. If $j \notin \mathcal{A}_k$ then $t_{ij} = \hat{t}_{ij} = (e - c) / 3 - d_{ij}$.

Definition 2. (World FTA) If the network g is connected, then there is a *world FTA*.

Definition 3. (Only-regional FTA, complete-regional FTA and extra-regional FTA) If there is a component for which all elements are in the same region, $i, j \in \mathcal{A}_k \subseteq R_k$, then \mathcal{A}_k is an *only-regional FTA*; if $\mathcal{A}_k = R_k$ and the network g_{R_k} is connected, then we say there is a *complete-regional FTA*; if $\mathcal{A}_k \subsetneq R_k$ then \mathcal{A}_k is

¹¹This notation emphasizes that for i and j to be linked there can either be a link from i to j , or from j to i or both. This is sometimes referred to as a non-directed link. As shall become clear, we need to make a distinction between directed and non-directed links. Whilst non-directed links determine which countries are members of an agreement, directed links determine the sponsor of the agreement. See Section 3.2 for further details.

an *extra-regional FTA*.

Definition 1 gives FTA membership a definition in graph notation, and a convenient graphical representation. Definition 2 then says that if all countries are in the same component then there is a world FTA. If, on the other hand, all members of a component are in the same region then the component is an only-regional FTA and if all countries in a region are in the same FTA then we say there is a complete-regional FTA. Finally, if FTA membership spans regions then the FTA is said to be extra-regional. A country that has no links with other countries is said to be in its own *singleton component*.

4.3.2 The Sponsor of an FTA

For a trade agreement to come about, we will say that it must have a sponsor. A sponsor must pay a *sponsorship cost* for coordinating an FTA. As mentioned in the introduction a sponsor is the country that meets the cost of bringing all other country representatives to the negotiating table in order to make the agreement for that period.

When a country sponsors an agreement between a group of countries not already in an agreement, it must pay the cost of bringing them all to the table. Formally, let the linear function $\kappa_i(z)$, which is increasing in z , give the cost of sponsoring an agreement with z countries. Let the function $\kappa_i(z)$ be the same for all i . So country i pays a cost $\kappa_i(1)$ for each link that it forms. The sponsorship cost for an agreement with countries $\mathcal{A}_j = \{j_1, \dots, j_l\}$ is $\kappa_i(|\mathcal{A}_j|)$.¹²

What happens if the set of countries $\mathcal{A}_j = \{j_1, \dots, j_l\}$ with whom Country i proposes to sponsor an FTA are themselves already in an FTA? Then we will say that Country i only has to pay the cost of a single link $\kappa_i(1)$; its own cost of

¹²Note that sponsorship costs do not depend on which country is being linked to. It is the same whether the country is in the same region or in a different region. It will become clear from the analysis below that if sponsorship of an agreement with more distant nations were more costly, this would reinforce the incentive to form regional only agreements.

reaching their table. If Country i is itself already in an FTA, entailing the set of countries $\mathcal{A}_k = \{i_1, \dots, i, \dots, i_l\}$, then we will say that Country i *acts as a delegate for* \mathcal{A}_k if it sponsors an FTA with \mathcal{A}_j ; that is, Country i proposes an FTA between all members of \mathcal{A}_j and \mathcal{A}_k and that again the sponsorship cost is just $\kappa_i(1)$; the cost of Country i reaching the table of \mathcal{A}_j .¹³

The fact that Country i must act as a delegate for its partners in \mathcal{A}_k is a strong simplifying assumption. In principle it makes world free trade more likely. If, from a given configuration of FTAs, a member of an existing FTA finds it worth sponsoring an agreement with another FTA then it will automatically bring all its existing FTA partners into the other FTA. Thus, the assumption effectively causes two FTAs to merge if any one member of one FTA joins another FTA. If regional FTA formation can be demonstrated under this assumption, then it seems even more likely to emerge in more realistic (but also more difficult to analyze) scenarios, in which members of a given FTA are allowed to ‘go it alone’ in approaching another FTA. Countries i and j in a given FTA could each sponsor agreements with two other separate FTAs, potentially resulting in much more complex equilibrium paths.

Nevertheless, it is conjectured that allowing Country i to ‘go it alone’ and join another FTA makes no difference to equilibrium of the present model. If there is no incentive for two blocks to merge then there would be no incentive for an individual country to deviate by joining an FTA in a different region; the sponsorship cost in both cases is the same and the production-trade gains to each member from the merger of two FTAs are greater. Admittedly, this suggests that the present model dramatically simplifies the actual incentives behind trade agreement formation.

¹³The term ‘coming to the table’ is used as a metaphor for participating in negotiations to form an FTA. Clearly, it involves more than just the cost of physically getting to the meeting point where the agreement is discussed. It also includes, at a minimum, the briefing costs of knowing what would be entailed by the agreement and the opportunity costs of politicians and officials of attending the meeting.

When Bhagwati, Greenaway and Panagariya (1998) represent the set of preferential trade agreements that actually existed in the mid-1990s using a network of the kind formalized here, they aptly refer to the resulting diagram as a “spaghetti bowl”.

Because a country pays a sponsorship fee for each agreement that it proposes, we will want a way of keeping track of the total amount that each country pays in sponsorship fees. To this end, define $\eta_i^\sigma(g) = |\{k \in \mathcal{N} \mid g_{ik} = 1\}|$ as the number of countries with which i maintains direct links. Then the total sponsorship cost paid by a country is given by $\kappa_i(\eta_i^\sigma(g))$ or $\eta_i^\sigma(g) \kappa_i(1)$.

As well as wishing to calculate the costs of FTA formation to each country, we will also want to calculate the benefits. Because in general these vary across regions, we will need to distinguish between the number of members in each. Define $\eta_i(g) = \left| \left\{ i \in R_k, j \in R_k \mid j \xrightarrow{g} i \right\} \cup \{i\} \right|$ $i \neq j$ as the number of countries in the same region as Country i and on the same path. Recall that D_i contains the set of distinct distances d_k of other countries from Country i . Define $\eta_i^{d_k}(g) = \left| \left\{ i \in R_k, j \in R_j \mid d_{ij} = d_k, j \xrightarrow{g} i \right\} \right|$, $i \neq j$, as the number of other countries $j \neq i$ at a distance $d_{ij} = d_k > 0$ on the same path as Country i . Let $\left\{ \eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_k}(g), \dots, \eta_i^{d_z}(g) \right\} = H_i(g)$ be the complete set of *membership variables* $\eta_i^{d_k}(g)$.

In the three region model, where each region is assumed to be at the vertex of an equilateral triangle, we have a particularly simple representation. All countries not in the same region as Country i are at the same distance away. So if $i \in R_k$ then $d_{ij} = d > 0$ for all $j \notin R_k$ and a single scalar which we can call $\eta_i^d(g)$ gives the total number of countries not in R_k with which Country i is linked.

The process of agreement formation will be much easier to formalize if we know that any agreement which some Country i proposes to sponsor will be accepted by all the proposed partners.¹⁴

¹⁴If this were not known ex ante, then we would need to model an explicit procedure by which the agreement were formed. This would entail a sponsor announcing the countries with which

Proposition 3. *Assume A1. If Country i proposes to sponsor an FTA with a set of other countries $\mathcal{A}_j = \{j_1, \dots, j_l\}$, then all countries in the set \mathcal{A}_j would obtain positive production-trade payoffs from the proposed FTA and would therefore accept. This holds whether or not the set of countries $\mathcal{A}_j = \{j_1, \dots, j_l\}$ are themselves already in an FTA and whether or not Country i is itself already in an FTA.*

This result is intuitively obvious. By Proposition 2, the production-trade gains to an FTA are always positive. So if some country proposes to sponsor an FTA then all of the proposed members will always accept. Therefore the proposal of an agreement is synonymous with its formation.

The following result shows that the network structure gives rise to an ordinary coalition structure of the form $\mathcal{C} = \{\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_m\}$.

Lemma 2. *An FTA structure is a partition of the set of countries \mathcal{N} : $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$; $\cup_{i \in \mathcal{C}} \mathcal{A}_i = \mathcal{N}$.*

Thus, an FTA structure is like the *cooperation structure* modelled by Myerson (1977). This follows from the simplifying assumption made above that any country acts as a delegate for its partners, if it is already in an FTA, when it proposes to join another FTA. We have a conventional coalition structure rather than a network. But as we shall see, the network terminology of link formation is helpful because it enables us to model equilibrium very conveniently in the manner of a non-cooperative network.

4.3.3 Payoffs to FTA Formation

The production-trade payoffs w_i , given by the function (4.8), will now be adapted for use as a payoff function in an FTA formation game.

Lemma 3. *Let $i \in \mathcal{A}_i$. Let x_{ij} be given by (4.7) and let the set of model*

it would like to sponsor an agreement, and then each proposed member accepting or rejecting the agreement in turn. Such a procedure has been formalized for coalition formation by Bloch (1996).

parameters be represented by γ . Then the production-trade payoff function (4.8) can be expressed in the form

$$w_i = w \left(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_k}(g), \dots, \eta_i^{d_z}(g); \gamma \right),$$

or equivalently

$$w_i = w(H_i(g); \gamma)$$

To gain greater insight into this result, look at the expanded form of the payoff function $w(H_i(g); \gamma)$:

$$\begin{aligned} & w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g), \dots; \gamma) \\ = & \eta_i(g) \left(\frac{3}{8} (e - c)^2 \right) \\ & + (n - |\mathcal{A}_i|) \left(\frac{5}{18} (e - c)^2 \right) \\ & + \eta_i^{d_1}(g) \left(\frac{1}{8} (3(e - c) + d_1)(e - c - d_1) \right) \\ & \vdots \\ & + \eta_i^{d_z}(g) \left(\frac{1}{8} (3(e - c) + d_z)(e - c - d_z) \right). \end{aligned}$$

The expression is parametric except for the membership variables $\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g)$ because tariffs have been substituted for, using either the optimal tariff formula (Proposition 1) where no agreement exists, or zero tariffs where an agreement exists.¹⁵

The first line shows the payoff to production and trade with FTA members that are in the same region, the number of which is given by $\eta_i(g)$. No distance parameter appears on this line because countries in the same region are assumed to

¹⁵Look at the proof of Lemma 3 to see how the terms in $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g), \dots; \gamma)$ are derived.

have the same location. The second line gives the payoff to production and trade with all countries *not* in Country i 's FTA. Notice that when optimal tariffs are in place, the volume of trade is exactly the same between all non-members of the FTA, regardless of their distance. The remaining lines measure the production-trade gains from FTA members at all distances $d_k \in D_i$.¹⁶

Having derived a convenient short-hand to write w_i in terms of regional FTA membership, it is now possible to evaluate the gains to Country i from FTA formation with other regional members and countries from outside the region. The following result shows that as long as optimal tariffs are positive, then an increase in membership of Country i 's FTA carries production-trade gains to Country i . The result also shows that more production-trade gains are derived the closer are the new members.

Let $\Delta\eta_i(g)$ denote a unit increase in $\eta_i(g)$ from any level and let $\Delta\eta_i^{d_k}(g)$ denote a unit increase in $\eta_i^{d_k}(g)$ from any level. Let $\Delta w/\Delta\eta_i(g)$ and $\Delta w/\Delta\eta_i^{d_k}(g)$ measure the impact on $w(H_i(g); \gamma)$ of a unit increase in $\eta_i(g)$ and $\eta_i^{d_k}(g)$ respectively.

Lemma 4. *Assume A1.*

(i) *The terms $\Delta w/\Delta\eta_i(g)$ and $\Delta w/\Delta\eta_i^{d_k}(g)$ are positive and constant, (where d_k is any element of the set D_i) and are independent of $\eta_i(g)$ and $\eta_i^{d_k}(g)$.*

(ii) *Let d_j be the smallest element of D_i , let $d_l \in D_i$ be the largest element, and let $d_k \in D_i$ be any other element such that $d_j < d_k < d_l$. Then*

$$\Delta w/\Delta\eta_i(g) > \Delta w/\Delta\eta_i^{d_j}(g) > \Delta w/\Delta\eta_i^{d_k}(g) > \Delta w/\Delta\eta_i^{d_l}(g) \geq 0.$$

¹⁶To see why no distance parameters appear in the second line showing production-trade gains with non-members of the FTA, use the expression for the optimal tariff (Proposition 1) in the expression for output (4.7) and notice that the distance parameter cancels.

If optimal tariffs are removed and tariffs are set to zero then the distance parameters appear. This explains why the distance parameters do appear on the remaining lines.

Lemma 4 extends Proposition 2 from bilateral to multilateral agreements. As mentioned earlier, due to the fact that there is no trade diversion, and that the firm in each country has a monopoly over the good that it exports, the production-trade gains to trade reaped by each citizen-producer can be calculated as a fixed multiple of the number of countries in the agreement.¹⁷

Assumption A1 ensures that optimal tariffs are positive. Part (i) of Lemma 4 establishes that $\Delta w / \Delta \eta_i(g)$ and $\Delta w / \Delta \eta_i^{d_k}(g)$ are constant and do not depend on the initial levels of $\eta_i(g)$ and $\eta_i^{d_k}(g)$. This is convenient because it means that the production-trade benefit of changes in FTA membership can be evaluated depending only on the distance between members. Consequently, the production-trade gains of *any* change in FTA membership, regional or non-regional, can be captured using the notation $\Delta w / \Delta \eta_i(g)$ and $\Delta w / \Delta \eta_i^{d_k}(g)$. For example, the effect of an increase in $\eta_i(g)$ from y^1 to y^2 is given by $(y^2 - y^1) \Delta w / \Delta \eta_i(g)$.

Part (ii) of Lemma 4 shows that production-trade gains of trade block expansion are greater for closer countries. To understand why, look at the payoff function $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g), \dots; \gamma)$; notice that the production-trade payoff from regional members, $\frac{3}{8}(e - c)^2$, is greater than the production-trade payoff from countries where there is no agreement, $\frac{5}{18}(e - c)^2$. The proof then shows that for any two countries joining an FTA, the production-trade payoff lies between these two levels. The intuition is straight-forward. Because tariffs between closer countries are higher, their removal brings about a relatively large increase in production-trade gains. Formally, this follows from Proposition 1, which shows optimal tariffs to be declining in distance, and Proposition 2(ii), which shows in turn that the production-trade gains to a bilateral agreement are declining in distance. As long as signing an FTA entails removal of positive tariffs, it must yield a positive gain. But this is ensured by Assumption A1, which guarantees that optimal tariffs are positive.

¹⁷In a more complex model, the degree of trade diversion would vary according to the specific FTA structure being considered.

Three Regions on an Equilateral Triangle Again

As mentioned above, when using the three region model only two parameters are needed to describe FTA membership from the point of view of Country i . These parameters are $\eta_i(g)$ and $\eta_i^d(g)$, which give regional and non-regional membership respectively. Nothing more complicated than the 3-region model, based on an equilateral triangle, is needed to motivate the tendency to form regional agreements, and the results will be based on this simplified special case.¹⁸

4.3.4 FTA Formation with Sponsorship Costs

Now that the production-trade payoffs of FTA formation have been determined, these can be used in the payoffs of a noncooperative network formation game. An overall payoff function will be specified in which the increased likelihood of being voted for as a result of the production-trade payoffs from FTA formation can be balanced against the sponsorship costs.

The parameters in the vector γ are held constant throughout, so from now on the function $w(H_i(g); \gamma)$ will be written $w(H_i(g))$. Using this form for production trade payoffs, define each country's *overall payoff function* $\Psi_i : \mathcal{G} \rightarrow \mathbb{R}$ as follows:

$$\begin{aligned} \Psi_i(g) &= \psi(H_i(g), \eta_i^\sigma(g)) \\ &= w(H_i(g)) - \kappa_i(\eta_i^\sigma(g)) \end{aligned} \tag{4.9}$$

The function (4.9) has been transformed so that the term $w(H_i(g))$ enters with no coefficient. More generally, we would expect this term to carry a coefficient which translates the increase in welfare from FTA formation into the increased likelihood of being voted for. Through the transformation, this effect is incorporated into the function $\kappa_i(\cdot)$.

When optimal tariffs are positive (A1), $w(\cdot)$ is increasing in $\eta_i(g)$, and weakly increasing in $\eta_i^{dk}(g) \in H_i(g)$ (Lemma 4). So for the purposes of the analysis

¹⁸Results will be generalized further in a subsequent version.

$(H_i(g), \eta_i^\sigma(g))$ is increasing in $\eta_i(g)$ and weakly increasing in $\eta_i^{d^*}(g)$.¹⁹ The function $\kappa_i(\cdot)$ is a linear function of $\eta_i^\sigma(g)$, the scalar measuring the number of FTAs sponsored by Country i . As already pointed out above, sponsorship costs are invariant to the distance between members.

4.3.5 FTAs in Equilibrium as Nash Networks

With payoffs to the network formation game now specified, we can define equilibrium as a Nash network (Bala and Goyal 2000). Given a network $g \in \mathcal{G}$, let g_{-i} denote the network obtained when all of Country i 's links are removed. Then the network g can be written as $g = g_i \oplus g_{-i}$, where \oplus denotes that g is formed as the union of the links in g_i and g_{-i} . The strategy g_i is a *best response* of Country i to g_{-i} if there does not exist a strategy g' for which

$$\Psi_i(g'_i \oplus g_{-i}) \geq \Psi_i(g_i \oplus g_{-i}) \text{ for all } g'_i \in \mathcal{G}_i.$$

The set of all Country i 's best responses to g_{-i} is denoted $BR_i(g_{-i})$. A network $g = (g_1, \dots, g_n)$ is a *Nash network* if $g_i \in BR_i(g_{-i})$ for each i .

This definition of equilibrium is a straightforward application of the standard notion of Nash equilibrium to a noncooperative network setting. A network is in a state of equilibrium if none of the agents, countries in the setting of this present paper, has an incentive to deviate. In the present setting, deviation would entail a country breaking a link by withdrawing its sponsorship of an FTA.

4.3.6 The Dynamics of Regionalism

This process is based on naive best response dynamics. The *FTA formation game* is assumed to last for three periods $t = 0, 1, 2$. The process is initialized with the empty network at $t = 0$. The FTA formation game is assumed to be repeated in

¹⁹Asymmetries in the value of links across players have also been considered in network formation models by Myerson (1980) and Slikker and van den Nouweland (2000) among others.

periods $t = 1, 2$. Within each period, the sequence of events is as follows. Each country observes the FTAs (described by the network g) of the previous period. Then, using (4.9) each country simultaneously chooses a *set of proposed partners*; Country i chooses \mathcal{P}_i , Country j chooses \mathcal{P}_j and so on; each \mathcal{P}_i contains the set of countries with which Country i would like to sponsor an agreement in the current period. Then each of the intersections $\cap_{i \in \mathcal{N}} \mathcal{P}_i$ forms a *proposed FTA*. A sponsor for each proposed FTA is then picked at random from its members. Payoffs that would result if the FTA were formed are calculated, where sponsorship costs are conditional upon the FTA structure of the previous period. If the payoff given by (4.9) is negative for the sponsor, then that country is allowed to refuse sponsorship, in which case another sponsor is picked at random, and so on.²⁰ If the payoff to all sponsors is negative then no FTA is formed. Then (conditional on the FTA structure of the previous period) sponsorship costs are paid by all sponsors, FTAs are formed and tariffs are set. Finally, production and trade take place for given tariffs and output is consumed. There is a *subgame perfect equilibrium* of the FTA formation game if there is a Nash network in period $t = 1, 2$.

4.4 The 3-Region FTA in Equilibrium

This section uses the simple 3-region model to present the main results of the paper. Nothing more complex than the 3-region model is needed to show why trade blocks may be regional. So let $|\mathcal{C}| = 3$. Countries are located at the vertices of an equilateral triangle. Then for all countries not sharing the same region,

²⁰This is a shortcut taken to ensure immediate convergence to a Nash network. It circumvents the collective action problem of forming such an agreement described by Olsen (1965). Other papers in the international trade literature simply assume that the group is small enough to overcome the collective action problem; see for example Grossman and Helpman (1994). On the other hand, to get around this problem a full dynamic process of convergence on a non-cooperative basis to a Nash network is modelled by Bala and Goyal (2000) and could be included here as well.

$i \in R_i, j \in R_j, i \neq j$, the distance between them is given by the same parameter, $d_{ij} = d > 0$, which measures the length of the sides of the triangle. Then the variable $\eta_i^d(g)$ measures the total number of non-regional members in Country i 's FTA. (The variable $\eta_i(g)$ gives the number of regional members as before.)

4.4.1 Production-Trade Payoffs, Sponsorship Costs and Overall Payoffs in a Network Game

One of the main advantages of using a 3-region model is that it keeps the overall payoff function as simple as possible. The overall payoff function takes the form

$$\psi(H_i(g), \eta_i^\sigma(g)) = \psi(\eta_i(g), \eta_i^d(g), \eta_i^\sigma(g)).$$

Equilibrium analysis will centre on showing network configurations from which there is no incentive to deviate. So we will want a method of examining the change in payoffs to all possible strategic alternatives that are available to a country.

To develop such a method, let $\Delta\eta_i^\sigma(g)$ denote a unit increase of $\eta_i^\sigma(g)$. This provides convenient notation to help evaluate the change in the overall payoff to the sponsorship of any given agreement. For example, suppose that Country i has already sponsored z agreements with countries; formally Country i has sponsored z links with other countries. And through these agreements Country i is in an FTA with y^1 other regional countries and y_d^1 countries outside the region. Then the payoff to the sponsorship of an additional agreement, which will enlarge the FTA to include $y^2 > y^1$ countries from the region and $y_d^2 > y_d^1$ from outside the region is given by

$$\begin{aligned} & \psi(y^2, y_d^2, z + \Delta\eta_i^\sigma(g)) - \psi(y^1, y_d^1, z) \\ &= (y^2 - y^1) \Delta w / \Delta\eta_i(g) + (y_d^2 - y_d^1) \Delta w / \Delta\eta_i^d(g) - \kappa_i(\Delta\eta_i^\sigma(g)). \end{aligned}$$

The left hand side takes the difference between overall payoffs under the two network structures. The first term on the right hand side shows the production-trade

gains to an increase in regional members of the FTA. The second term shows the production-trade gains to an increase in non-regional members. The third term shows the sponsorship costs of setting up the additional agreement. Taken together, these terms show how the production-trade gains balance and the reciprocal expected gains in votes balance against the sponsorship costs of an agreement.²¹

Recall from Lemma 4 that the production-trade payoffs to a regional FTA are higher than to a non-regional FTA. From what we have just seen, it is easy to envisage sponsorship costs at a level where regional agreements of a given size are worthwhile but non-regional agreements are not. The following result formalizes this idea by looking at sponsorship costs across a range of levels and their implications for the incentive to sponsor regional and non-regional FTAs. It is important to keep in mind when looking at this result, however, that it evaluates the incentive to form an FTA where none are pre-existing. As we shall see later, the production-trade payoffs to form new FTAs from existing ones are greater than the incentive to get FTAs off the ground in the first place.

Lemma 5. *Assume A1. Assume that in period $t = 0$ the network g is empty.*

(i) *Let the production-trade payoff to a bilateral agreement with a country in the same region be lower than the sponsorship cost. If there are no existing FTAs then no FTA is worth sponsoring. Formally, $\Delta w / \Delta \eta_i(g) < \kappa(1) \Rightarrow \psi(y_1, y_1^d, (y_1 + y_1^d - 1)) < \psi(1, 0, 0)$ for $1 < r \leq y^1$, $0 \leq y_d^1 \leq n - r$.*

(ii) *Let the production-trade payoff to a bilateral agreement with a country in a different region be higher than the sponsorship cost. Even if there are no FTAs, then the payoff to sponsorship of a world FTA is higher than the payoff to sponsorship of any other FTA. Formally, $\Delta w / \Delta \eta_i^d(g) > \kappa(1) \Rightarrow \psi(r, n - r, (n - 1)) \geq \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$, for $r \geq y^1 \geq 1$, $n - r \geq y_d^1 \geq 0$, holding with strict inequality if and only if $y^1 < r$ and/or $y_d^1 < n - r$.*

(iii) *Let the production-trade payoff to a bilateral agreement with a country in*

²¹In the following, we will just refer to “production-trade gains” and drop the reference to “the reciprocal expected gains in votes”.

the same region be higher than the sponsorship cost. But let the production-trade payoff to a bilateral agreement with a country in a different region be lower than (or equal to) the sponsorship cost. If there are no existing FTAs then sponsorship of a complete-regional FTA yields a higher payoff than sponsorship of any other agreement. Formally, $\Delta w / \Delta \eta_i^d(g) < \kappa(1) < \Delta w / \Delta \eta_i(g) \Rightarrow \psi(r, 0, r-1) > \psi(r, n-r, n-1)$ and $\psi(r, 0, r-1) > \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$, for $r > y_1^1 \geq 1$, $n-r \geq y_d^1 \geq 1$.

In a situation where there are no FTAs already existing, Lemma 5 shows the FTA structure that will yield the highest payoff from sponsorship. If the production-trade payoffs of a bilateral agreement are lower than the sponsorship cost even for an agreement with regional neighbors, then no country has an incentive to sponsor an FTA (Lemma 5(i)). If the production-trade payoffs are higher than the sponsorship cost of an FTA with a country in another region then an a world FTA will be worth sponsoring (Lemma 5(ii)).

It is when costs are at an intermediate level that the incentives show scope for regionalism. In part (iii) of Lemma 5 it is assumed that the costs of production-trade payoffs of sponsoring an FTA with a country in the same region are above the sponsorship costs. Therefore, it is immediately clear that it will be worth sponsoring an FTA with regional neighbors. But sponsorship costs are above the production-trade payoffs of an FTA with countries outside the region. So from a situation where a country did sponsor an extra-regional FTA, it would gain more from withdrawing its sponsorship of an agreement with those more distant nations than from the production-trade gains of maintaining it. In this situation, the sponsorship costs lie between the relatively large gains from removing higher mutual tariffs with close neighbors and the smaller gains from removing lower mutual tariffs with countries that are further away. An only-regional FTA is the only type of agreement that is worth sponsoring.

Lemma 5 focuses exclusively on the payoffs to a country when it is the sole sponsor of an FTA. It will be shown that in equilibrium any given FTA can only

have one sponsor. But first, to make analysis of the equilibrium path easier, it will be helpful to look at how the incentives to sponsor a world FTA change when starting not from a situation where there are no FTAs but from one where there are complete-regional FTAs already in existence. The incentives to sponsor an extra-regional FTA, given that a complete-regional FTA already exists, are analyzed in the next result.

Lemma 6. *Assume A1.*

Let the production-trade payoff to a bilateral agreement with a single country in a different region be lower than the sponsorship cost. Assume that a complete-regional agreement exists in every region $R_k \in P$. An extra-regional FTA (formed with a single link) is worth sponsoring if the production-trade payoffs to an FTA with more than one country in a different region are higher than the sponsorship cost of a bilateral agreement. (If no FTA already exists in another region R_j then an extra-regional agreement is not worth sponsoring.)

Formally, $y_1^d \Delta w / \Delta \eta_i^d(g) > \kappa(1) > \Delta w / \Delta \eta_i^d(g) \Rightarrow \psi(y_1, ay_1^d, y_1 + a - 1) > \psi(y_1, 0, y_1 - 1)$, for $r \geq y^1 \geq 1$, $r \geq y^1 > 1$, $a \geq 1$

If there are enough other countries from another region already in an FTA then the production-trade benefits may overcome the sponsorship costs, even though these costs are too high to make an agreement with a single other country in that region worthwhile. This can happen because it is assumed that negotiating with an existing FTA incurs only the sponsorship cost of a bilateral agreement $\kappa(1)$. Recall that if Country i wants to negotiate membership with an existing FTA then it only has to pay the cost of bringing itself to their table. And there exists a range of $\kappa(1)$ for which the production-trade gains to a bilateral agreement with a country outside the region is less than $\kappa(1)$, but the production-trade gains to joining an existing FTA with more than one country are greater than $\kappa(1)$. The last part of the Lemma, shown in brackets, is a re-statement of Lemma 5(ii), to emphasize the contrasting outcomes depending on whether or not an FTA exists

in the other region.

4.4.2 Equilibrium Paths; Are FTAs Stepping Blocks or Stumbling Blocks?

This subsection takes its title from the famous question posed by Bhagwati (1992). In the way that it will be answered below, the question should in fact be posed as follows: ‘When are FTAs stepping stones and when are they stumbling blocks in the path to free trade?’ As argued in the introduction of this present paper, trade blocks in the real world are regional. In this light, the question is whether the regional blocks presently existing will ultimately promote world free trade.

The term regionalism usually describes a situation where countries in a region form a club or agreement, but where membership does not extend beyond regional boundaries. For a corresponding analytical definition that will be useful in the present context, let *regionalism* be a situation where all regions have a complete FTA but where there are no extra-regional FTAs; in the network g there is a connected subnetwork g_k for each $R_k \in \mathcal{C}$, but $g_{ij} = g_{ji} = 0$ for all $i \in R_i, j \in R_j, i \neq j$. The next proposition presents the main result of the paper.

Proposition 4. *Assume A1.*

(i) *If the production-trade payoff to a bilateral agreement with a country in the same region is lower than the sponsorship cost then on the equilibrium path no FTA will exist at any point in time $t = 1, 2$*

(ii) *If the production-trade payoff to a bilateral agreement with a country in a different region is higher than the sponsorship cost then on the equilibrium path there is world free trade at every point in time $t = 1, 2$*

(iii) *On the equilibrium path there is regionalism in the first period if the following conditions hold:*

(a) *production-trade payoff to a bilateral agreement with a country in the same region is higher than the sponsorship cost*

(b) *production-trade payoff to a bilateral agreement with a country in a different region is lower than the sponsorship cost.*

(iv) *(Regional trade blocks are stepping blocks to free trade) On the equilibrium path there is regionalism in period $t = 1$ followed by world free trade in period $t = 2$ if the following conditions hold:*

(a) *the production-trade payoff to an extra-regional FTA with all countries in a different region is higher than the sponsorship cost (of a bilateral agreement);*

(b) *the production-trade payoff to a bilateral agreement with a single country in a different region is lower than the sponsorship cost.*

(c) *the production-trade payoff to a bilateral agreement with a single country in the same region is higher than the sponsorship cost.*

If the production-trade payoff to an extra-regional FTA with all countries in a different region is lower than the sponsorship cost then on the equilibrium path there is regionalism at every point in time.

Proposition 4 shows that the equilibrium path to free trade may indeed exhibit a period of regionalism followed by free trade (Proposition 4(iv)), presenting an encouraging answer to Bhagwati's question. Indeed, this is the most interesting possibility and the only one that is not immediately obvious. Let us briefly review the other outcomes before looking in more detail at Proposition 4(iv).

Obviously, if the sponsorship costs are prohibitive of even an FTA between close regional neighbors then none will be sponsored at all (Proposition 4(i)). On the other hand, with sponsorship costs sufficiently low there will be a move straight to free trade, bypassing regionalism altogether (Proposition 4(ii)). With costs at an intermediate level it is worth sponsoring an FTA between regional partners but not with countries that are further away, because gains to an FTA with regional partners are higher than with more distant nations.

If regionalism prevails in the first period, the question of whether the process stalls at regionalism thereafter or whether the process proceeds to free trade depends once again on the level of sponsorship costs. If the sponsorship cost (of

a bilateral agreement) is greater than the production-trade benefit of a bilateral agreement with a single country outside the region, but less than the production-trade benefit of joining a complete regional agreement outside the region then countries will wait until they see regional agreements formed elsewhere before proposing to join them. Of course, if the sponsorship cost (of a bilateral agreement) is higher than the production-trade benefit of joining a complete-regional agreement in another region then the process stalls at regionalism.

Finally, note that if one country sponsors an agreement with another region then free trade results immediately because of the assumption that any given country acts as a delegate for all of its regional partners when it negotiates to join an FTA with another region.

Recall that the cost structure assumed here has two bases, one in practice and one in theory. In practice, trade negotiators often report that it is much easier (and therefore cheaper) to negotiate an agreement with a block of countries than with countries as individuals (CREDIT 1998). In theory, Bala and Goyal assume that there is a single link cost of linking to an existing network; it is not necessary to pay a link cost of linking to each country individually.

4.5 Conclusions

The main purpose of this chapter has been to show that regionalism can arise in equilibrium. That is, countries may choose to form regional trade agreements rather than move all the way to free trade. Politicians balance the increased likelihood that they will be voted for if they coordinate an FTA against the coordination cost itself. The increased likelihood of being voted for results from conventional welfare gains to trade due to formation of the FTA. By working out optimal tariffs with transport costs, it is shown that trade based gains to an agreement with countries of the same region are higher than gains to an agreement involving (distant) countries from different regions. On the other hand, the costs of bringing

politicians together from different nations in order to coordinate an agreement are assumed to be proportional to the number of countries involved, and not dependent upon which countries they come from. Therefore, a regional FTA may be worth coordinating whilst one involving countries from outside the region may not.

In addition, Chapter 3 also makes a contribution to the debate on whether trade blocks are building blocks or stumbling blocks in the freeing of trade multilaterally. It shows that a period of regionalism may be necessary in order for free trade to take place. This result was based on the idea that it is cheaper to sponsor an agreement with an existing FTA than with all its individual members; an idea expressed in practice by trade negotiators and formalized in the literature on noncooperative network formation.

There are a number of extensions to this work that suggest themselves immediately. One straightforward extension to appear in the next version of this paper is to present the results of this paper for any number of regions. A more substantive extension would be to give greater attention to the negotiation process in the model. In the chapter as it stands, the model of agreement formation is crude. Explicit models of agreement formation, have been constructed by Busch and Wen (1995) and Furusawa and Wen (2002) and it would be interesting to extend these to a regional setting of the kind described in Chapter 4.

Also, it would be nice to improve upon the ‘naive’ best response dynamics assumed in Chapter 4; the basis for the evolution of regional agreements, possibly towards free trade. It would be far more satisfactory to undertake such analysis, and establish whether or not trade blocks are building blocks, in a network where agents are far sighted as in Page and Wooders (2002) for example.

4.6 Appendix

Proof of Lemma 1. First rearrange the expression for CS_i as follows:

$$\begin{aligned}
 CS_i &= \sum_{j \in \mathcal{N}} \frac{1}{2} (e - p_{ij}) x_{ij} \\
 &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - x_{ij} - c_{ij}) x_{ij} \\
 &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c_{ij}) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 \\
 &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c - t_{ij} - d_{ij}) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2
 \end{aligned}$$

where the second line follows by (4.5), and the fourth line follows by (4.3).

Using this, the expressions for TR_i , DR_i , and (4.6) in (4.8) yield

$$\begin{aligned}
 w &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c - t_{ij} - d_{ij}) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 \\
 &\quad + x_{ii}^2 + \sum_{j \in \mathcal{N}} t_{ij} x_{ij} + \sum_{j \in \mathcal{N}} d_{ij} x_{ij} + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2 \\
 &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} t_{ij} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} d_{ij} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 \\
 &\quad + x_{ii}^2 + \sum_{j \in \mathcal{N}} t_{ij} x_{ij} + \sum_{j \in \mathcal{N}} d_{ij} x_{ij} + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2 \\
 &= \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} t_{ij} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} d_{ij} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 \\
 &\quad + x_{ii}^2 + \sum_{j \in \mathcal{N}} t_{ij} x_{ij} + \sum_{j \in \mathcal{N}} d_{ij} x_{ij} + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2
 \end{aligned}$$

Now rearranging terms,

$$\begin{aligned}
 w &= \sum_{j \in \mathcal{N}} (e - c) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c - t_{ij} - d_{ij}) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 \\
 &\quad + x_{ii}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2 \\
 &= \sum_{j \in \mathcal{N}} (e - c) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} (e - c_{ij}) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 + x_{ii}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2 \\
 &= \sum_{j \in \mathcal{N}} (e - c) x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} (p_{ij} - c_{ij}) x_{ij} - \sum_{j \in \mathcal{N}} x_{ij}^2 + x_{ii}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2 \\
 &= (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 - \sum_{j \in \mathcal{N}/\{i\}} x_{ij}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2.
 \end{aligned}$$

where the third line uses (4.3). \square

Proof of Proposition 1. Using the expression for w obtained in Lemma 1, the government of Country i solves the following problem to set the optimal tariff on imports from Country j ;

$$\max_{t_{ij}} w = (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{3}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 + \sum_{j \in \mathcal{N}} x_{ji}^2.$$

Because, by (4.7), x_{ji} is not a function of t_{ij} (the tariff set by Country i does not affect production in other countries), the derivatives with respect to t_{ij} of all the terms under the last summation are equal to zero. Given that $-(x_{ij})^2$ is concave, and by (4.7) x_{ij} is a linear function of t_{ij} , the objective function w is concave in t_{ij} , all $j \in \mathcal{N}$. So there must exist a unique solution for \hat{t}_{ij} . Write the first order condition as

$$\frac{dw}{dt_{ij}} = (e - c) \frac{dx_{ij}}{dt_{ij}} - 3 \sum_{j \in \mathcal{N}} (x_{ij}) \frac{dx_{ij}}{dt_{ij}} = 0.$$

Then, using the fact that $dx_{ij}/dt_{ij} = -1/2$, simplifying and rearranging obtains $\hat{t}_{ij} = (e - c)/3 - d_{ij}$. It is immediate that if $d < (e - c)/3$ then $\hat{t}_{ij} > 0$. \square

Proof of Proposition 2. (i) Country i 's production-trade payoff is given by

$$w_i = (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2 - \sum_{j \in \mathcal{N}/\{i\}} x_{ij}^2 + \sum_{j \in \mathcal{N}/\{i\}} x_{ji}^2.$$

By (4.7), if $d_{ij} = d_{ji}$ and $t_{ij} = t_{ji}$ then $x_{ij} = x_{ji}$ and the last two terms cancel, leaving

$$w_i = (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2.$$

Using (4.7),

$$w_i = \sum_{j \in \mathcal{N}} \frac{1}{8} (3(e - c) + t_{ij} + d_{ij}) (e - c - t_{ij} - d_{ij}).$$

Now let countries i and j form an FTA. Using $t_{ij} = \hat{t}_{ij}$ for the pre-agreement tariff, and $t_{ij} = 0$ for the post agreement tariff in w_i , take discrete differences to work out the welfare gain:

$$\Delta w_i = \frac{1}{72} (7(e - c) + 3d_{ij}) ((e - c) - 3d_{ij}).$$

Under the assumptions that $e - c > 0$, and $t_{ij} = \hat{t}_{ij} = (e - c)/3 - d_{ij} > 0$, so $\Delta w_i > 0$.

(ii) Immediate by differentiation. \square

Proof of Proposition 3. By Definition 1, if $j_1 \in \mathcal{A}_j$ and $j_2 \in \mathcal{A}_j$ then mutual tariffs are set at $t_{j_1 j_2} = t_{j_2 j_1} = 0$. By Proposition 2, each country in \mathcal{A}_j gains $\Delta w_i > 0$ for each other country in the agreement. Moreover, each country in \mathcal{A}_j pays no sponsorship cost. Therefore, it is in the interest of each country in the set \mathcal{A}_j to accept membership of the FTA proposed by Country i . \square

Proof of Lemma 2. Suppose not. $\cup_{i \in \mathcal{C}} \mathcal{A}_i = \mathcal{N}$ is trivial (recall that countries not in an FTA are singleton components.) To see that $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$, suppose not. Suppose that $i \in \mathcal{A}_i$ and $i \in \mathcal{A}_j$. This may be the case for one of the following reasons. Either i proposed to sponsor \mathcal{A}_i and \mathcal{A}_j . But in that case there must be a path between i and all members of $\mathcal{A}_i \setminus \{i\}$ and i and all members of $\mathcal{A}_j \setminus \{i\}$. But by Definition 1 all members of \mathcal{A}_i and \mathcal{A}_j must be in the same FTA; a contradiction. Or i was already in one FTA, without loss assume \mathcal{A}_i , and proposed to sponsor an agreement with the members of \mathcal{A}_j . But then, by assumption, if i 's proposal were accepted, all members of \mathcal{A}_i must have joined \mathcal{A}_j at the same time;

a contradiction. Finally, suppose that i was already in one agreement, assume \mathcal{A}_i , but came into \mathcal{A}_j as a result of an agreement proposed by another country. But then if i entered \mathcal{A}_j in this way, then so must all $\mathcal{A}_i \setminus \{i\}$; a contradiction. \square

Proof of Lemma 3. For convenience, define the following piece of notation. Let $\bar{g}_{ij} = \max \{g_{ij}, g_{ji}\}$. Note that, by (4.7), $t_{ij} = t_{ji}$ and $d_{ij} = d_{ji}$, it is the case that $x_{ij} = x_{ji}$ for all $i, j \in \mathcal{N}$ (independent of whether $g_{ij} = 0$ or $g_{ij} = 1$). Consequently, the function w can be written in the form

$$w_i = (e - c) \sum_{j \in \mathcal{N}} x_{ij} - \frac{1}{2} \sum_{j \in \mathcal{N}} x_{ij}^2.$$

Let $x_{ij}^{d_k}(\bar{g}_{ij})$ represent (4.7) where the superscript d_k denotes that Country j is at distance $d_k > 0$ from Country i ; $i \in R_i$, $j \in R_j$, $i \neq j$. We substitute for (4.7) explicitly in the step after this. But to see how the structure of the new function arises, it is helpful to note the following intermediate step. Recall that $g_{ij} \in \{0, 1\}$, where t_{ij} is set optimally according to Proposition 1 if $\bar{g}_{ij} = 0$ and free trade is adopted if and only if $\bar{g}_{ij} = 1$. As $x_{ij}^{d_k}(\bar{g}_{ij})$ depends only on \bar{g}_{ij} and d_k , the function w_i can be partitioned accordingly:

$$\begin{aligned} w_i &= w\left(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_k}(g), \dots, \eta_i^{d_z}(g); \gamma\right) \\ &= \eta_i(g) \left((e - c) x_{ij}(1) - \frac{1}{2} x_{ij}(1)^2 \right) \\ &\quad + (r - \eta_i(g)) \left((e - c) x_{ij}(0) - \frac{1}{2} x_{ij}(0)^2 \right) \\ &\quad + \eta_i^{d_1}(g) \left((e - c) x_{ij}^{d_1}(1) - \frac{1}{2} x_{ij}^{d_1}(1)^2 \right) \\ &\quad + (\delta_{1i} - \eta_i^{d_1}(g)) \left((e - c) x_{ij}^{d_1}(0) - \frac{1}{2} x_{ij}^{d_1}(0)^2 \right) \\ &\quad \vdots \\ &\quad + \eta_i^{d_z}(g) \left((e - c) x_{ij}^{d_z}(1) - \frac{1}{2} x_{ij}^{d_z}(1)^2 \right) \\ &\quad + (\delta_{zi} - \eta_i^{d_z}(g)) \left((e - c) x_{ij}^{d_z}(0) - \frac{1}{2} x_{ij}^{d_z}(0)^2 \right) \end{aligned}$$

where the absence of a superscript in the terms $x_{ij}(1)$ and $x_{ij}(0)$ denotes that $i, j \in R_k$.

Now substitute explicitly for (4.7). First note that if $i, j \in R_k$ then $d_{ij} = 0$. Also, if $g_{ij} = 1$ then $t_{ij} = 0$ and, by (4.7), $x_{ij}(1) = (e - c)/2$. If $\bar{g}_{ij} = 0$ then $t_{ij} = (e - c)/3$ and so $x_{ij}(0) = (e - c)/3$. Analogously, by (4.7), $x_{ij}^{d_k}(1) = (e - c - d_k)/2$. And, by Proposition 1, use $\hat{t}_{ij} = \frac{e-c}{3} - d_k$ and (4.7) to obtain $x_{ij}^{d_k}(0) = (e - c)/3$. Making these substitutions, we can rewrite the function $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g), \dots; \gamma)$ as

$$\begin{aligned} & w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g), \dots; \gamma) \\ = & \eta_i(g) \left(\frac{3}{8} (e - c)^2 \right) \\ & + (r - \eta_i(g)) \left(\frac{5}{18} (e - c)^2 \right) \\ & + \eta_i^{d_1}(g) \left(\frac{1}{8} (3(e - c) + d_1)(e - c - d_1) \right) \\ & + (\delta_{k1} - \eta_i^{d_1}(g)) \left(\frac{5}{18} (e - c)^2 \right) \\ & \vdots \\ & + \eta_i^{d_z}(g) \left(\frac{1}{8} (3(e - c) + d_z)(e - c - d_z) \right) \\ & + (\delta_{kz} - \eta_i^{d_z}(g)) \left(\frac{5}{18} (e - c)^2 \right) \end{aligned}$$

Now, using the facts that $r + \sum_{d_k \in D_i} \delta_{ki} = |\mathcal{N}|$ and $\eta_i(g) + \sum_{d_k \in D_i} \eta_i^{d_k}(g) = |\mathcal{A}_i|$ we can simplify further by writing

$$\begin{aligned} & w(\eta_i(\bar{g}), \eta_i^{d_1}(\bar{g}), \dots, \eta_i^{d_z}(\bar{g}), \dots; \gamma) \\ = & \eta_i(\bar{g}) \left(\frac{3}{8} (e - c)^2 \right) \\ & + (|\mathcal{N}| - |\mathcal{A}_i|) \left(\frac{5}{18} (e - c)^2 \right) \\ & + \eta_i^{d_1}(\bar{g}) \left(\frac{1}{8} (3(e - c) + d_1)(e - c - d_1) \right) \\ & \vdots \\ & + \eta_i^{d_z}(\bar{g}) \left(\frac{1}{8} (3(e - c) + d_z)(e - c - d_z) \right). \end{aligned}$$

□.

Proof of Lemma 4: (i) To show that $\Delta w / \Delta \eta_i(g)$ is constant, begin by noting that although $\eta_i(g)$ is a discrete variable, the function $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g); \gamma)$ is continuous in $\eta_i(g)$. Treating $\eta_i(g)$ as a continuous variable in a compact set, it is possible to calculate the derivative of $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g); \gamma)$ with respect to $\eta_i(g)$;

$$\frac{\partial w}{\partial \eta_i(g)} = \frac{7}{72} (e - c)^2$$

As the expression for $\partial w / \partial \eta_i(g)$ is parametric, the effect on w of a discrete change in η_i is given by

$$\Delta w = \left(\frac{7}{72} (e - c)^2 \right) \Delta \eta_i(g)$$

This holds at *any* $\eta_i(g)$, as required. As $(e - c) > 0$ by assumption, $\Delta w / \Delta \eta_i(g) > 0$.

To show that $\Delta w / \Delta \eta_i^{d_k}(g)$ is constant, follow the same procedure. Calculate the derivative of $w(\eta_i(g), \eta_i^{d_1}(g), \dots, \eta_i^{d_z}(g); \gamma)$ with respect to $\eta_i^{d_k}(g)$:

$$\frac{\partial w}{\partial \eta_i^{d_k}(g)} = \frac{1}{72} (7(e - c) + 3d_k)(e - c - 3d_k)$$

As the expression for $\partial w / \partial \eta_i^{d_k}(g)$ is parametric, the effect on w of a discrete change in $\eta_i^{d_k}$ is given by

$$\Delta w = \frac{1}{72} (7(e - c) + 3d_k)(e - c - 3d_k) \Delta \eta_i^{d_k}(g)$$

Again, this holds at *any* $\eta_i(g)$ as required. By A1, $(e - c) > 3d_k$ and therefore $\Delta w / \Delta \eta_i^{d_k}(g) > 0$.

(ii) Show that $\Delta w / \Delta \eta_i(g) > \Delta w / \Delta \eta_i^{d_j}(g) > \Delta w / \Delta \eta_i^{d_k}(g) > \Delta w / \Delta \eta_i^{d_l}(g) \geq 0$. First establish that $\Delta w / \Delta \eta_i(g) > \Delta w / \Delta \eta_i^{d_j}(g)$. From (i) we know that in general

$$\frac{\Delta w}{\Delta \eta_i^{d_k}(g)} = \frac{1}{72} (7(e - c) + 3d_k)(e - c - 3d_k)$$

where d_k is any element of D_i . Expanding the brackets,

$$\frac{\Delta w}{\Delta \eta_i^{d_k}(g)} = \frac{1}{72} (7(e-c)^2 - 18(e-c)d_k - 9d_k^2)$$

Notice that $\Delta w/\Delta \eta_i^{d_k}(g)$ is declining in d_k , attaining its maximum for $d_k = 0$. So $\Delta w/\Delta \eta_i(g) > \Delta w/\Delta \eta_i^{d_k}(g)$ for all $d_k \in D_i$ and, in particular, $\Delta w/\Delta \eta_i(g) > \Delta w/\Delta \eta_i^{d_j}(g)$.

Next, establish that $\Delta w/\Delta \eta_i^{d_j}(g) > \Delta w/\Delta \eta_i^{d_k}(g) > \Delta w/\Delta \eta_i^{d_l}(g)$. But this follows immediately by the fact that $\Delta w/\Delta \eta_i^{d_k}(g)$ is declining in d_k , and that by assumption $d_j < d_k < d_l$.

Finally, it must be established that $\Delta w/\Delta \eta_i^{d_l}(g) \geq 0$. The root for $\Delta w/\Delta \eta_i^{d_k}(g) = 0$ is $d_k = (e-c)/3$. To see this, use $d_k = (e-c)/3$ in $\Delta w/\Delta \eta_i^{d_k}(g)$ to obtain

$$\frac{\Delta w}{\Delta \eta_i^{d_k}(g)} = \frac{1}{72} \left(7(e-c)^2 - \frac{18}{3}(e-c)^2 - 9 \left(\frac{e-c}{3} \right)^2 \right) = 0.$$

But by A1, $0 < d_k \leq (e-c)/3$. The result follows. \square

Proof of Lemma 5. By A1, $\Delta w/\Delta \eta_i(g) > \Delta w/\Delta \eta_i^d(g) \geq 0$ (Lemma 4).

(i) Assume that initially g is the empty network and suppose to the contrary that there does exist an FTA that is worth sponsoring. For this to be the case the overall payoff to sponsoring such an agreement must be higher than autarchy. Then there exist values of y^1 and y_d^1 , where $r \geq y^1 > 1$, $n-r \geq y_d^1 \geq 0$, for which $\psi(y_1, y_1^d, (y_1 + y_1^d - 1)) > \psi(1, 0, 0)$. This implies

$$\begin{aligned} & \psi(y_1, y_1^d, (y_1 + y_1^d - 1)) - \psi(1, 0, 0) \\ &= (y_1 - 1) \Delta w/\Delta \eta_i(g) + y_1^d \Delta w/\Delta \eta_i^d(g) - \kappa(y_1 + y_1^d - 1) > 0. \end{aligned}$$

But

$$\begin{aligned} & (y_1 - 1) \Delta w/\Delta \eta_i(g) + y_1^d \Delta w/\Delta \eta_i^d(g) - \kappa(y_1 + y_1^d - 1) \\ &= (y_1 - 1) \Delta w/\Delta \eta_i(g) + y_1^d \Delta w/\Delta \eta_i^d(g) - (y_1 + y_1^d - 1) \kappa(1) \\ &< (y_1 - 1) \Delta w/\Delta \eta_i(g) + y_1^d \Delta w/\Delta \eta_i(g) - (y_1 + y_1^d - 1) \kappa(1) \\ &= (y_1 + y_1^d - 1) (\Delta w/\Delta \eta_i(g) - \kappa(1)), \end{aligned}$$

and by assumption $\Delta w/\Delta \eta_i(g) < \kappa(1)$ so $(y_1 + y_1^d - 1)(\Delta w/\Delta \eta_i(g) - \kappa(1)) < 0$; contradiction.

(ii) Assume that initially g is the empty network and suppose to the contrary that sponsorship of some FTA other than the world FTA yields a higher payoff. Then $\psi(r, n - r, (n - 1)) < \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$ for all values of y^1 and y_d^1 , where $r \geq y^1 \geq 1$, $n - r \geq y_d^1 \geq 0$. This implies

$$\begin{aligned} & \psi(r, n - r, (n - 1)) - \psi(y_1, y_1^d, (y_1 + y_1^d - 1)) \\ &= (r - y_1) \Delta w/\Delta \eta_i(g) + (n - r - y_1^d) \Delta w/\Delta \eta_i^d(g) - \kappa(n - y_1 - y_1^d) < 0. \end{aligned}$$

But

$$\begin{aligned} & (r - y_1) \Delta w/\Delta \eta_i(g) + (n - r - y_1^d) \Delta w/\Delta \eta_i^d(g) - \kappa(n - y_1 - y_1^d) \\ &= (r - y_1) \Delta w/\Delta \eta_i(g) + (n - r - y_1^d) \Delta w/\Delta \eta_i^d(g) - (n - y_1 - y_1^d) \kappa(1) \\ &> (r - y_1) \Delta w/\Delta \eta_i^d(g) + (n - r - y_1^d) \Delta w/\Delta \eta_i^d(g) - (n - y_1 - y_1^d) \kappa(1) \\ &= (n - y_1 - y_1^d) (\Delta w/\Delta \eta_i^d(g) - \kappa(1)), \end{aligned}$$

and by assumption $\Delta w/\Delta \eta_i^d(g) > \kappa(1)$ so $(n - y_1 - y_1^d) (\Delta w/\Delta \eta_i^d(g) - \kappa(1)) \geq 0$; contradiction. Clearly, if $y^1 = r$ and $y_d^1 = n - r$ then $(n - y_1 - y_1^d) (\Delta w/\Delta \eta_i^d(g) - \kappa(1)) = 0$ and $\psi(r, n - r, (n - 1)) = \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$. But if $y^1 < r$ and/or $y_d^1 < n - r$ then $(n - y_1 - y_1^d) (\Delta w/\Delta \eta_i^d(g) - \kappa(1)) > 0$. The result follows.

(ii) Assume that initially g is the empty network and suppose to the contrary that sponsorship of some FTA other than the complete-regional FTA yields a higher payoff. Then $\psi(r, 0, (r - 1)) < \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$ for some values of y^1 and y_d^1 , where $r > y^1 \geq 1$, and $n - r \geq y_d^1 \geq 0$. This implies

$$\begin{aligned} & \psi(r, 0, (r - 1)) - \psi(y_1, y_1^d, (y_1 + y_1^d - 1)) \\ &= (r - y_1) \Delta w/\Delta \eta_i(g) + (n - r - y_1^d) \Delta w/\Delta \eta_i^d(g) - \kappa(n - y_1 - y_1^d) < 0. \end{aligned}$$

But

$$\begin{aligned}
 & (r - y_1) \Delta w / \Delta \eta_i(g) + (n - r - y_1^d) \Delta w / \Delta \eta_i^d(g) - \kappa(n - y_1 - y_1^d) \\
 = & (r - y_1) \Delta w / \Delta \eta_i(g) + (n - r - y_1^d) \Delta w / \Delta \eta_i^d(g) - (n - y_1 - y_1^d) \kappa(1) \\
 > & (r - y_1) \Delta w / \Delta \eta_i^d(g) + (n - r - y_1^d) \Delta w / \Delta \eta_i^d(g) - (n - y_1 - y_1^d) \kappa(1) \\
 = & (n - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)),
 \end{aligned}$$

and by assumption $\Delta w / \Delta \eta_i^d(g) > \kappa(1)$ so $(n - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) \geq 0$; contradiction. Clearly, if $y^1 = r$ and $y_d^1 = n - r$ then $(n - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) = 0$ and $\psi(r, n - r, (n - 1)) = \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$. But if $y^1 < r$ and/or $y_d^1 < n - r$ then $(n - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) > 0$. The result follows.

(iii) Assume that initially g is the empty network and suppose to the contrary that sponsorship of some FTA other than the complete-regional FTA yields a higher payoff. Then either $\psi(r, 0, (r - 1)) < \psi(r, n - r, (n - 1))$ or $\psi(r, 0, (r - 1)) < \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$ for all values of y^1 and y_d^1 , where $r > y^1 > 1$, and $n - r \geq y_d^1 \geq 0$. But the first inequality implies

$$\begin{aligned}
 & \psi(r, 0, (r - 1)) - \psi(r, n - r, (n - 1)) \\
 = & -(n - r) \Delta w / \Delta \eta_i^d(g) + \kappa((n - r)) \\
 = & -(n - r) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) < 0.
 \end{aligned}$$

and by assumption $\Delta w / \Delta \eta_i^d(g) < \kappa(1)$ and so $-(n - r) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) > 0$; contradiction.

The second inequality implies

$$\begin{aligned}
 & \psi(r, 0, (r - 1)) - \psi(y_1, y_1^d, (y_1 + y_1^d - 1)) \\
 = & (r - y_1) \Delta w / \Delta \eta_i(g) - y_1^d \Delta w / \Delta \eta_i^d(g) - \kappa(r - 1 - y_1 - y_1^d + 1) < 0.
 \end{aligned}$$

But

$$\begin{aligned}
 & (r - y_1) \Delta w / \Delta \eta_i(g) - y_1^d \Delta w / \Delta \eta_i^d(g) - \kappa(r - 1 - y_1 - y_1^d + 1) \\
 = & (r - y_1) \Delta w / \Delta \eta_i(g) - y_1^d \Delta w / \Delta \eta_i^d(g) - (r - y_1 - y_1^d) \kappa(1) \\
 > & (r - y_1) \Delta w / \Delta \eta_i^d(g) - y_1^d \Delta w / \Delta \eta_i^d(g) - (r - y_1 - y_1^d) \kappa(1) \\
 = & (r - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)),
 \end{aligned}$$

and by assumption $\Delta w / \Delta \eta_i^d(g) < \kappa(1)$ so $(r - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) > 0$ for $r < y_1 + y_1^d$. Now $(r - y_1 - y_1^d) (\Delta w / \Delta \eta_i^d(g) - \kappa(1)) < 0$ for $r > y_1 + y_1^d$. But in addition

$$\begin{aligned}
 & (r - y_1) \Delta w / \Delta \eta_i(g) - y_1^d \Delta w / \Delta \eta_i^d(g) - (r - y_1 - y_1^d) \kappa(1) \\
 < & (r - y_1) \Delta w / \Delta \eta_i(g) - y_1^d \Delta w / \Delta \eta_i(g) - (r - y_1 - y_1^d) \kappa(1) \\
 = & (r - y_1 - y_1^d) (\Delta w / \Delta \eta_i(g) - \kappa(1)),
 \end{aligned}$$

and by assumption $\kappa(1) < \Delta w / \Delta \eta_i(g)$, so $(r - y_1 - y_1^d) (\Delta w / \Delta \eta_i(g) - \kappa(1)) > 0$ for $r > y_1 + y_1^d$. \square

Proof of Lemma 6. Let there be a regional agreement of size $r \geq y_1^d > 1$ in region R_j ($i \in R_i$, $i \neq j$). The proof is in two parts. (i) Show that if $y_1^d \Delta w / \Delta \eta_i^d(g) > \kappa(1) > \Delta w / \Delta \eta_i^d(g)$ then Country i does find it worth sponsoring an extra-regional agreement with the FTA in R_j . (ii) Show that this does not hold if $\kappa(1) > y_1^d \Delta w / \Delta \eta_i^d(g)$.

(i) Assume $y_1^d \Delta w / \Delta \eta_i^d(g) > \kappa(1) > \Delta w / \Delta \eta_i^d(g)$. Suppose to the contrary that an extra-regional agreement with the FTA in R_j is not worth sponsoring. This implies

$$\begin{aligned}
 & \psi(y_1, y_1^d, y_1) - \psi(y_1, 0, y_1 - 1) \\
 = & y_1^d \Delta w / \Delta \eta_i^d(g) - \kappa(1) < 0
 \end{aligned}$$

But by assumption $y_1^d \Delta w / \Delta \eta_i^d(g) > \kappa(1)$; contradiction.

(ii) Now assume $\kappa(1) > y_1^d \Delta w / \Delta \eta_i^d(g)$ in order to see that Country i does not find it worth sponsoring an extra-regional agreement with the FTA in R_j .

Suppose to the contrary that such an agreement is worth sponsoring. This implies $\psi(y_1, y_1^d, y_1) - \psi(y_1, 0, y_1 - 1) = y_1^d \Delta w / \Delta \eta_i^d(g) - \kappa(1) > 0$; contradiction. \square

Proof of Proposition 4.

First observe that under the 3-region structure, existence of a nonempty Nash network depends upon finding a member of each FTA that will be its sponsor. By the symmetry of the 3-region structure, if $j \in \mathcal{P}_i$ then $i \in \mathcal{P}_j$ and so $\mathcal{P}_i = \mathcal{P}_j$. Therefore, for an FTA to exist it must be shown that any given member would obtain a positive payoff from sponsoring it. This will be shown for each case below.

(i) Suppose to the contrary that there exists a period in which at least one FTA is sponsored. Let $t = s$ be the first period in which at least one FTA is sponsored. Then by the assumption that the network is empty at $t = 0$, the network must be empty at $t = s - 1$. Taking as given the empty network g at $t = s - 1$, the payoff to sponsoring an agreement with y_1 countries in the same region and y_1^d countries outside the region is $\psi(y_1, y_1^d, (y_1 + y_1^d - 1))$. By assumption, production-trade payoffs and sponsorship costs are in the same relation as in Lemma 5(i). But by Lemma 5(i), $\psi(1, 0, 0) > \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$ and therefore any country sponsoring an agreement could gain by deleting all its links; withdrawing sponsorship. So the network must be empty at $t = s$ as well. As $t = s$ is any period $t \geq 1$, the network g must be empty at every period $t \geq 1$.

(ii) Suppose to the contrary that there exists a period $t \geq 1$ in which there is not a world FTA. Let $t = s$ be a period in which the Nash network g is either empty or not connected. By definition of equilibrium, there must exist a world FTA at $t = 1$. By assumption, production-trade payoffs and sponsorship costs are in the same relation as in Lemma 5(ii). Then by Lemma 5(ii), $\psi(r, n - r, (n - 1)) \geq \psi(y_1, y_1^d, (y_1 + y_1^d - 1))$, for $r \geq y_1 \geq 1$, $n - r \geq y_1^d \geq 0$ holding with strict inequality if and only if $y_1 < r$ and/or $y_1^d < n - r$. So the empty network cannot be Nash; Country i would receive a payoff $\Psi_i = \psi(1, 0, 0) < \psi(r, n - r, (n - 1))$ and has an incentive to deviate by forming links with all other countries. Similarly, if Country i sponsors an FTA that is not a world FTA, then it can increase its

payoff by forming links, again contradicting Nash.

Given that the Nash network g is connected at $t = 1$, then it must be connected at $t = 2$. If not, then the sponsor of the agreement, Country i , must have deleted some or all of its links. But deviating in this way would yield a lower payoff than maintaining all links; $\psi(y_1, y_1^d, (y_1 + y_1^d - 1)) < \psi(r, n - r, (n - 1))$, $r > y^1 \geq 1$, $n - r > y_d^1 \geq 0$, contradicting equilibrium. By induction, taking as given a world FTA in period $t = s - 1$, it is a best response for any country to maintain sponsorship of the world FTA at $t = s$. As $t = s$ is any period $t > 1$, and as the network g is connected at $t = 1$, it must be connected at every period $t \geq 1$.

(iii) By definition, there is regionalism in the network g if there is a connected subnetwork g_k for each $R_k \in \mathcal{C}$, but $g_{ij} = g_{ji} = 0$ for all $i \in R_i$, $j \in R_j$, $i \neq j$. Suppose to the contrary that the Nash network g does not exhibit regionalism. There are two (mutually inclusive) possibilities. One is that the Nash network g contains links $g_{ij} = 1$ or $g_{ji} = 1$ for some $i \in R_i$, $j \in R_j$, $i \neq j$. The other is that the subnetwork g_k is not connected for some $R_k \in \mathcal{C}$. Contradictions for these two possibilities are found in turn.

Suppose to the contrary that at $t = 1$ the Nash network g contains links $g_{ij} = 1$ or $g_{ji} = 1$ for some $i \in R_i$, $j \in R_j$, $i \neq j$. By assumption, production-trade payoffs and sponsorship costs are in the same relation as in Lemma 5(iii). Then by Lemma 5(iii), $\psi(r, 0, r - 1) > \psi(r, y_1^d, (r + y_1^d - 1))$, for $r > y^1 \geq 1$, $n - r \geq y_d^1 \geq 1$. Therefore, if Country $i \in R_i$ sponsors any links of the form $g_{ij} = 1$ with $j \in R_j$, $i \neq j$, then it can gain by deleting them, so the network g cannot be Nash. By the same argument, Country i has an incentive to break links if it sponsors a world FTA.

Now suppose that in the Nash network g of period $t = 1$, Country i sponsors an FTA that is not a complete-regional FTA; that is where $y^1 < r$. But then again by Lemma 5(iii) $\psi(r, 0, r - 1) > \psi(y_1, y_1^d, (r + y_1^d - 1))$, for $r > y^1 \geq 1$, $n - r > y_d^1 \geq 0$. Therefore, Country i could gain by linking to the other countries j for which $i, j \in R_i$, so the network g cannot be Nash. It follows that for each

$R_k \in \mathcal{C}$ the subnetwork g_k must be connected. By definition of equilibrium, only one country sponsors the complete-regional FTA. If not then a second sponsor could withdraw from sponsorship, gaining the sponsorship cost and not losing any production-trade payoffs, contradicting Nash.

(iv) Conditions (b) and (c) are exactly as in (iii) so from (iii) we know that there must be regionalism at $t = 1$. Take the network from period $t = 1$ as given, where the subnetworks g_k are connected for the elements of all $R_k \in \mathcal{C}$. Suppose to the contrary that at $t = 2$ the Nash network g is not connected. By assumption, production-trade payoffs and sponsorship costs are in the same relation as in Lemma 5(iii). If at $t = 2$ any subnetwork g_k is not connected then by Lemma 5(iii) there is an incentive to deviate by forming links to other countries within R_k so g cannot be Nash. Moreover, condition (a) implies that production-trade payoffs and sponsorship costs are in the same relation as in Lemma 6. So by Lemma 6, if there does not exist a link between Country $i \in R_i$ and $j \in R_j$, $i \neq j$, then Country i could gain by forming a link with a country in another region, $\psi(r, ay_1^d, r + a - 1) > \psi(r, 0, r - 1)$, $0 < ay_1^d \leq (n - r)$, $a \geq 1$, contradicting Nash.

Finally, suppose that condition (a) does not hold, so that the payoff from linking to an FTA in another region is not greater than the sponsorship cost. Then the Nash network g cannot be complete, because if Country $i \in R_i$ sponsors any links to countries $j \in R_j$, then it could gain by breaking those links. However, given that conditions (b) and (c) continue to hold, and given regionalism at $t = 1$, there is no incentive for any country to deviate at $t = 2$. By (iii), any sponsor of a regional agreement could not gain by deleting links. So there is regionalism in at $t = 1$. Under these same conditions, given regionalism at $t = s - 1$, there must be regionalism at $t = s$. So when (a) fails to hold there must be regionalism at all points on the equilibrium path. \square

Chapter 5

Conclusions

This dissertation has examined situations where there are conflicts of interest between governments in policy setting, focusing in particular on tax and trade policy. There is a large literature in each of these areas. Yet there appear to be reasons to believe that existing theory does not give us a complete understanding of the tax and trade policy interactions we have been observing.

The first chapter of the dissertation adapts Hotelling's model of competition between firms to think about tax competition between governments. Hotelling's Principle of Minimum Differentiation is applied in the context of tax competition and shown to be invalid. There may be excessive differentiation in which one government overprovides the amenity and sets taxes too high whilst the other government underprovides the amenity and sets taxes too low. Not only may there be excessive differentiation but, in addition, equilibrium may fail to exist. Non-existence of equilibrium is interpreted as a metaphor for intense tax competition.

Chapter 1 represents the first occasion on which, to our knowledge, Hotelling's model and the possible nonexistence of equilibrium have been adapted to think about amenities and taxation competition. The model seems to provide a way of understanding the patterns of taxation that have been observed in Europe, for example, where the core has not been forced to reduce taxes significantly towards the lower rates set in the periphery, despite ever more integrated capital markets.

Also in Chapter 1, perfect tax discrimination was analyzed to examine the extent to which it provided a solution to the equilibrium existence issue raised by Hotelling (uniform) tax competition. We find that allowing governments to discriminate perfectly in setting taxes only partially resolves the equilibrium existence problem. There is a larger range of values for which the cost of amenity mismatch supports an equilibrium. But even under perfect tax discrimination, if the cost of amenity mismatch is relatively high then tax competition is so intense that the system does not settle down to an equilibrium. Moreover, when equilibrium does exist, the efficiency problem is exacerbated. A race to the bottom occurs in which neither government offers amenities at a positive level in equilibrium.

Chapters 2 and 3 switch focus from problems of international tax competition to problems with forming trade agreements. Chapter 2 examined the multilateral trade liberalization process. In Chapter 2, the fact that trade liberalization has proceeded in a number of rounds over the post-war period was explained in terms of the institutional structure imposed on trade liberalization by the GATT rules themselves. Attention focused particularly on a ‘withdrawal of equivalent concessions,’ (WEC) and tariff bindings.

The first main result of Chapter 2 was that was that the WEC rule does facilitate trade liberalization but that free trade certainly cannot be reached, no matter how little countries discount the future. A standard repeated game allows trade partners to implement the worst (credible) punishment against a deviant. In general, the WEC rule makes such severe punishments illegal. By outlawing a class of severe punishments, the WEC rule compromises efficiency. The second main result was that if punishments are constrained by the WEC rule *and* initial deviations are bound, then on the equilibrium path, trade liberalization must take place over a number of periods or ‘rounds’. Intuitively, these rules effectively limit countries’ ability to withdraw cooperation. As withdrawals of cooperation are limited, only the promise of future cooperation is available to make an agreement today self-enforcing. As a result cooperation, in this instance over tariff liberalization, cannot

be given away all at once. It must be given up gradually.

Chapter 3 proposed an answer to the question of why trade agreements tend to be regional. It argued that politicians balance the increased likelihood that they will be voted for if they coordinate an FTA against the coordination cost itself. The increased likelihood of being voted for results from conventional welfare gains to trade due to formation of the FTA. By working out optimal tariffs with transport costs, it is shown that trade based gains to an agreement with countries of the same region are higher than gains to an agreement involving (distant) countries from different regions. On the other hand, the costs of bringing politicians together from different nations in order to coordinate an agreement are assumed to be proportional to the number of countries involved, and not dependent upon which countries they come from. Therefore, a regional FTA may be worth coordinating whilst one involving countries from outside the region may not.

In addition, Chapter 3 also made a contribution to the debate on whether trade blocks are building blocks or stumbling blocks in the freeing of trade multilaterally. It showed that a period of regionalism may be necessary in order for free trade to take place. This result was based on the idea that it is cheaper to sponsor an agreement with an existing FTA than with all its individual members; an idea expressed in practice by trade negotiators and formalized in the literature on noncooperative network formation.

5.1 Four Broad Themes

Four broad themes emerge from the research. The first is that the forces of competition over taxation between governments may not simply push all governments in the same direction; towards efficiency or underprovision of amenities. There are good reasons to suppose that governments may be pushed in opposing directions in order to soften the degree of competition that they face. By making their jurisdictions different from others, by offering a particular level or type of amenity,

jurisdictions may effectively be able to reduce the extent to which they have to compete over conventional measures such as taxation. This thinking appears to provide stylized outcomes that are in some ways closer to what we actually observe. It also seems this thinking may extend to a range of other areas, including environmental policy, where it was previously believed that competition forced jurisdictions in a single direction.

The second broad theme is that in fact harmonization measures may be motivated not by inefficiency but by intense competition. It is impossible to comment on the efficiency of a situation where no equilibrium exists. But it seems fair to conclude that it is not a pleasant environment for government officials to work in. It may be the intensity of competition, more than the inefficiency that results, that may be motivating calls for tax harmonization.

The third broad theme is that cooperation may be gradual in any situation where there are restrictions on reversing levels of cooperation once they have been adopted. These restrictions need not be complete; partial irreversibility say in cooperation over investment or military matters will make cooperation necessarily gradual.

The fourth broad theme is that politicians may weight costs of forming agreements against different gains across different types of agent in deciding which agreements to form. Moreover, if the costs of linking to a network once it has been formed are the same as the costs of linking to a single player, then the results suggest that a set of small networks may expand into a bigger network subsequently.

5.2 Directions for Future Research

First, more work is needed to characterize situations where no equilibrium exists in the Hotelling tax competition model of Chapter 1. It appears to be an interesting idea that non-existence of equilibrium might represent a situation of intense competition. That is, from any given policy situation, there is always an incentive to

undercut another government. This appears to capture the difficulties that many governments complain of in a world where policy making is increasingly subject to external influences. Yet describing a situation of interest through non-existence of equilibrium is unsatisfactory in that it is not subject to further analysis; in particular, its efficiency properties cannot be analyzed. Therefore, more work is needed to capture situations of perpetual policy change through competition of this kind.

More work is also warranted on tax harmonization, given that it represents a way of escaping intense tax competition highlighted above. In the framework of Chapter 1 where governments are Leviathans, tax harmonization is likely to be bad for citizens in that governments could cooperate on tax setting to extract maximal rent from producers. A number of distributional issues are raised as well as issues for efficiency. But it is not possible even to begin to make comparisons in an environment where equilibrium does not exist under Hotelling tax competition. So further research that preserves the basic features of intense competition whilst enabling equilibrium and efficiency issues to be analyzed could be very useful.

Extending the ideas in Chapter 2, there are a number of useful ideas that could usefully be pursued in future work on gradualism in a policy environment. One is to extend the framework of partial irreversibility to other areas where cooperation is required, such as environmental policy or investment. Some work has already been undertaken in this area, by Chisik and Davies (2003) for example, under complete irreversibility. But the motivating forces are different from those of Lockwood and Thomas (2002) and further insights may be possible from extending the framework of partial irreversibility developed in Chapter 2 to other areas.

Perhaps more interesting still is the idea of using other modelling frameworks to probe in other ways the question of why negotiation of international treaties takes so long. One element that is missing from the model of Chapter 2 and may be an important factor in bringing about gradualism is that valuations over a trade agreement are private information. Bulow and Klemperer (1999) suggest a framework in which private valuations can be taken into account to understand

gradualism under the GATT.

There are a number of future directions suggested by Chapter 4 as well. One is to give greater attention to the negotiation process. In the chapter as it stands the model of agreement formation is crude. Explicit models of agreement formation, have been constructed by Busch and Wen (1995) and Furusawa and Wen (2002) and it would be interesting to extend these to a regional setting of the kind described in Chapter 4.

Also, it would be nice to improve upon the ‘naive’ best response dynamics assumed in Chapter 4; the basis for the evolution of regional agreements, possibly towards free trade. It would be far more satisfactory to undertake such analysis, and establish whether or not trade blocks are building blocks, in a network where agents are far sighted as in Page and Wooders (2002) for example.

Bibliography

- [1] Anderson, J.E. and H. van den Berg (1998); “Fiscal Decentralization and Government Size: An International Test for Leviathan Accounting for Unmeasured Economic Activity,” *International Tax and Public Finance* 5: 171-86.
- [2] Bagwell, K. and R. Staiger (1990); “A Theory of Managed Trade.” *American Economic Review*, 80, 779-795.
- [3] Bagwell, K. and R. Staiger (1999); “An Economic Theory of the GATT.” *American Economic Review*, 89(1) (March): 215-48.
- [4] Bala, V. and S. Goyal (2000); “A Noncooperative Model of Network Formation.” *Econometrica*, 68(5): 1181-1229.
- [5] Baldwin, R. and P. Krugman (2000); “Agglomeration, Integration and Tax Harmonization,” CEPR Discussion Paper no. 2630 (November).
- [6] Bergsten, F. (1975); *Toward A New International Economic Order*, Lexington Books, Lexington, MA.
- [7] Berliant, M. and F. Page (2001); “Income Taxes and the Provision of Public Goods: Existence of an Optimum,” *Econometrica*, 69(3): 771-84.
- [8] Besley, T. and M. Smart (2001); “Does Tax Competition Raise Voter Welfare?” LSE Typescript.

- [9] Bhagwati, J. (1988); *Protectionism*. MIT Press, Cambridge, MA.
- [10] Bhagwati, J. (1991); *The World Trading System at Risk*, Princeton University Press, Princeton.
- [11] Bhagwati, J. (1992); "Regionalism versus Multilateralism." *The World Economy*, 15: 535-555.
- [12] Bhagwati, J. (1993); "Regionalism and Multilateralism; an Overview," in de Melo and Panagariya (eds.) *New Dimensions in Regional Integration*, Cambridge University Press, Cambridge.
- [13] Bhagwati, J. D. Greenaway and A. Panagariya (1998); "Trading Preferentially: Theory and Policy," *Economic Journal*, (July) 108: 1128-1148.
- [14] Bhaskar, V. and T. To (2002); "Is Perfect Price Discrimination Really Efficient? An Analysis of Free Entry," *Rand Journal of Economics* forthcoming.
- [15] Black, D.A. and W. Hoyt (1989) "Bidding for Firms," *American Economic Review*, 79(5): 1249-1256.
- [16] Bloch, F. (1996); "Sequential Formation of Coalitions with Fixed Payoff Division and Externalities," *Games and Economic Behavior*, 14: 90-123.
- [17] Boadway, R. P. Pestieau and D.E. Wildasin (1989); "Non-cooperative Behavior and Efficient Provision of Public Goods," *Public Finance*, 44(1): 1-7
- [18] Boadway, R., K. Cuff and N. Marceau (2002); "Inter-jurisdictional Competition for Firms," *International Economic Review*, 43(5): 761-782.
- [19] Bond, E. (1999) "Multilateralism, Regionalism, and the Sustainability of 'Natural' Trading Blocs." Penn State University mimeograph.
- [20] Bond, E.B. and J. H. Park (2000); "Gradualism in Trade Agreements with Asymmetric Countries." To appear in *Review of Economic Studies*.

- [21] Bond, E. and L. Samuelson (1986); "Tax Holidays as Signals," *American Economic Review*, 76(4): 820-826.
- [22] Bond, E.B. and C. Syropoulos (1996); "Trading Blocs and the Sustainability of Inter-regional Cooperation." Chapter in M. Canzoneri, W. Ethier and V. Grilli (eds), *The New Transatlantic Economy*, pp. 118-141. Cambridge University Press, Cambridge, UK.
- [23] Bond, E.B., C. Syropoulos and A. Winters (2001); "Deepening of regional integration and multilateral trade agreements." *Journal of International Economics*, 53: 335-361.
- [24] Brainard, S.L. and T. Verdier (1994); "Lobbying and adjustment in declining industries," *European Economic Review*, 38: 586-595.
- [25] Brander, J. and B. Spencer (1984); "Tariff Protection and Imperfect competition," chapter in H. Kierkowski (ed.) *Monopolistic Competition and International Trade*, Oxford University Press, Oxford.
- [26] Brennan, G. and J.M. Buchanan, (1980); *The Power To Tax: Analytical Foundations of a Fiscal Constitution*, Cambridge University Press, Cambridge.
- [27] Busch, L.-A., and Q. Wen (1995); "Perfect Equilibria in a Negotiation Model," *Econometrica*, 63(3): 545-565.
- [28] Bucovetsky, S. (1991); "Asymmetric Tax Competition," *Journal of Urban Economics*, 30: 167-81.
- [29] Bulow, J. and P. Klemperer, (1999); "The Generalized War of Attrition," *American Economic Review*, 89(1): 175-89.

- [30] Burbidge, J. B., J. DePater, G.M. Myers and A. Sengupta (1994). "A Coalition Formation Approach to Equilibrium Federations and Trading Blocs," *American Economic Review* 87, 940-956.
- [31] Cassing, J.H. and A.L. Hillman, (1986); "Shifting comparative advantage and senescent industry collapse." *American Economic Review*, 76(3): 516-523.
- [32] Chisik, R. and R.B. Davies (2003); "Gradualism in Tax Treaties with Irreversible Foreign Direct Investment." *International Economic Review*, forthcoming.
- [33] Coase, R.H. (1960); "The Problem of Social Costs", *Journal of Law and Economics*.
- [34] Conconi, P. (2002); *Trade and the Environment*, PhD Thesis, University of Warwick.
- [35] Conley, J. and M. Wooders (1997); "Equivalence of the Core and Competitive Equilibrium in a Tiebout Economy with Crowding Types." *Journal of Urban Economics*, 41, 421-440.
- [36] Conley, J. and M. Wooders (2001); "Tiebout Economies with Differential Genetic Types and Endogenously Chosen Crowding Characteristics." *Journal of Economic Theory*, 98(2): 261-94 (June).
- [37] CREDIT, University of Nottingham (1998) "Study on the economic impact of introducing reciprocity in the trade relations between the EU and EAC countries", June.
- [38] Dam, K.W. (1970); *The GATT: Law and International Economic Organization*. Chicago University Press, Chicago.

- [39] Dasgupta, P. and E. Maskin (1986); "The Existence of Equilibrium in Discontinuous Economic Games I. Theory." *Review of Economic Studies* 53, 1-26.
- [40] d'Aspremont, C., J. Gabszewicz, and J. Thisse (1979); "On Hotelling's Stability in Competition," *Econometrica*, 17: 1145-1151.
- [41] DePater, J.A. and G.M. Myers, (1994); "Strategic Capital Tax Competition: A Pecuniary Externality and a Corrective Device," *Journal of Urban Economics* 36: 66-78.
- [42] Devereux, M. (1997) "Growth, specialization and trade liberalization." *International Economic Review*, 38(3): 565-585.
- [43] Devereux, M.P., R. Griffith and A. Klemm (2001); "Have Taxes on Mobile Capital Declined?" Institute for Fiscal Studies and University of Warwick, May 2001 unpublished manuscript.
- [44] Devereux, M., B. Lockwood and M. Redoano (2002); "Do Countries Compete over Corporate Tax Rates?" Warwick Economic Research Paper no. 642.
- [45] Dixit, A. (1987); "Strategic Aspects of Trade Policy." Chapter 9 in Bewley, T. F. *Advances in Economic Theory; Fifth World Congress*, Cambridge University Press, Cambridge.
- [46] Edwards, J. and M. Keen (1996); "Tax Competition and Leviathan." *European Economic Review* 40(1): 113-134.
- [47] Frankel, J., E. Stein and S. Wei (1995); "Trading Blocs and the Americas: The Natural, the Unnatural and the Super-natural." *Journal of Development Economics*, 47: 61-95.
- [48] Furusawa, T. and H. Konishi (2002) "Free Trade Networks." Yokohoma National University and Boston College mimeograph.

- [49] Furusawa, T. and L.-C. Lai (1999); "Adjustment costs and gradual trade liberalization." *Journal of International Economics*, 49: 333-361.
- [50] Furusawa, T. and Q. Wen (2002); "Disagreement Points in Trade Negotiations," *Journal of International Economics*, 57: 133-50.
- [51] General Agreement on Tariffs and Trade (GATT 1994); "Uruguay Round: Final Act" Marrakesh, 15 April. Photocopy.
- [52] Gordon, R. and J.D. Wilson (1986); "An Examination of Multijurisdictional Corporate Income Taxation under Formula Apportionment," *Econometrica*, 54(6): 1357-1373.
- [53] Gordon, R. and J.D. Wilson (1999); "Tax Structure and Government Behavior: Implications for Tax Policy," NBER Working Paper 7244.
- [54] Gordon, R. and J.D. Wilson (2002); "Expenditure Competition," *Journal of Public Economic Theory* forthcoming.
- [55] Goyal and Joshi (2000); "Bilateralism and Free Trade." Erasmus University and George Washington University mimeograph.
- [56] Grossman, G.M. and E. Helpman (1994) "Protection for Sale." *American Economic Review*, September, 84(4): 833-50.
- [57] Grossman, G.M. and E. Helpman (1995a) "Trade Wars and Trade Talks." *Journal of Political Economy*, August, 103(4): 675-708.
- [58] Grossman, G.M. and E. Helpman (1995b) "The Politics of Free Trade Agreements." *American Economic Review*, September, 85(4): 667-690.
- [59] Haaparanta, P. (1996); "Competition for Foreign Direct Investment," *Journal of Public Economics*, 63(1): 141-53.

- [60] Higgott, R. (1999); "Economics, Politics and (International) Political Economy: The Need for a Balanced Diet in an Era of Globalization." *New Political Economy*, 4(1): 23-36.
- [61] Hobbes, T. (1651); *Leviathan* reprinted subsequently in the Everyman Library series by J.M. Dent and Sons Ltd. 1914, 1962.
- [62] Hohaus, B., K. Konrad and M. Thum (1994); "Too much conformity? A Hotelling model of local public goods supply," *Economics Letters*, 44(3): 295-299.
- [63] Horwell, D.J. (1966); "Optimum Tariffs and Tariff Policy." *Review of Economic Studies*, 33, 147-158.
- [64] Hotelling, H. (1929); "Stability in Competition," *Economic Journal*, 39: 41-57.
- [65] Jackson, J. (1989); *The World Trading System*. MIT Press, Cambridge Massachusetts, 417 pages.
- [66] Jackson, M.O. and A. Wolinski, (1996) "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71: 44-74.
- [67] Johnson, H. G. (1953-54); "Optimum Tariffs and Retaliation." *Review of Economic Studies*, 21, 142-153.
- [68] Justman, M., J-F. Thisse and T. Van Ypersele (2001); "Fiscal competition and regional differentiation," CORE Discussion Paper No. 01/24.
- [69] Kemp, M. C., and H. J. Wan (1976); "An Elementary Proposition Concerning the Formation of Customs Unions", *Journal of International Economics* 6, 95-97.
- [70] Kennan, J., and R. Riezman (1990); "Optimal Tariff Equilibria with Customs Unions", *Canadian Journal of Economics* 90, 70-83.

- [71] Keohane, R. (1982); "The Demand for International Regimes." *International Organization* 36(2) (Spring) 325-355.
- [72] Keohane, R. (1984); *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press, Princeton, New Jersey.
- [73] King, I., R.P. McAfee and L. Welling (1993); "Industrial Blackmail: Dynamic Tax Competition and Public Investment," *Canadian Journal of Economics*, 26: 590-608.
- [74] Kose, M. A., and R. Riezman (1999). "Understanding the Welfare Implications of Preferential Trade Agreements", CSGR Working Paper 45/99, University of Warwick.
- [75] Kreps, D., P. Milgrom, D. Roberts and R. Wilson (1982); "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." *Journal of Economic Theory*, 27, 253-279.
- [76] Kreps, D. and J.A. Scheinkman (1983); "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *The Bell Journal of Economics*, 14(2): 326-337.
- [77] Kreps, D. and R. Wilson (1982); "Reputation and Imperfect Information." *Journal of Economic Theory*, 27, 253-279.
- [78] Krugman, P. (1991); "Is Bilateralism Bad?" in E. Helpman and J. Razin (eds.) *International Trade and Trade Policy*, MIT Press, Cambridge, MA.
- [79] Krugman, P. (1998); "What's New About the New Economic Geography," *Oxford Review of Economic Policy*, 14(2): 7-17.
- [80] Leamer, E. (1980) "Welfare computations and the optimal staging of tariff reductions in models with adjustment costs." *Journal of International Economics*, 10: 21-36.

- [81] Levy, P. (1997); "A Political-Economic Analysis of Free Trade Agreements." *American Economic Review*, 87(4): 506-19.
- [82] Lockwood, B. and J. Thomas (2002); "Gradualism and Irreversibility." *Review of Economic Studies* 69(2) no. 239: 339-356.
- [83] Lockwood, B. and K.Y. Wong (2000); "Specific and Ad Valorem Tariffs Are Not Equivalent in Trade Wars." *Journal of International Economics*, 52, 183-195.
- [84] Mayer, W. (1981); "Theoretical Considerations of Negotiated Tariff Adjustments." *Oxford Economic Papers*, 33, 135-153.
- [85] McLaren, J. (1997); "Size, Sunk Costs and Judge Bowker's Objection to Free Trade." *American Economic Review*, 87, 400-420.
- [86] Milgrom, P. and J. Roberts (1982); "Limit Pricing and Entry Under Incomplete Information," *Econometrica* 50: 443-460.
- [87] Mill, J. (1844); *Essays on Some Unsettled Questions of Political Economy*, Parker, London.
- [88] Mintz, J. and H. Tulkens, (1996); "Optimality Properties of Alternative Systems of Taxation of Foreign Capital Income." *Journal of Public Economics*, (June) 60(3): 373-99.
- [89] Mintz, J. and M. Smart (2001) "Income Shifting, Investment and Tax Competition: Theory and Evidence from Provincial Taxation in Canada," CESifo Working paper 554, August 2001, 17 pages.
- [90] Mussa, M. (1986) "The adjustment process and the timing of trade liberalization." Chapter 4 in A.M. Choski and D. Papageorgiou (eds.) *Economic Liberalization in Developing Countries*, Blackwell, Oxford.

- [91] Myerson, R. (1977); "Graphs and Cooperation in Games." *Mathematics of Operations Research* 2: 225-229.
- [92] Oates, W.E. (1972); *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.
- [93] Oates, W.E. (1985); "Searching for Leviathan: An Empirical Study," *American Economic Review*, 75: 748-757.
- [94] Oates, W.E. (1989); "Searching for Leviathan: A Reply and Some Further Reflections," *American Economic Review*, 79: 578-83.
- [95] Oates, W.E. and R.M. Schwab (1988); "Economic Competition Among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics*, 35: 333-354.
- [96] Olson, M. (1965). *The Logic of Collective Action*. Harvard University Press, Cambridge, MA.
- [97] Ottaviano, G.I.P. T. Tabuchi and J.-F. Thisse (2002); "Agglomeration and Trade: Revisited," *International Economic Review*, 43: 409-436.
- [98] Page, F. and M. Wooders (2002); "Networks and Farsighted Stability," Warwick Economic Research Papers no. 660, December 2002. Available on-line at <http://www.warwick.ac.uk/economics/research/twerps.html>.
- [99] Panagariya, A. (1998) "Do Transport Costs Justify Regional Preferential Trading Arrangements? No." *Weltwirtschaftliches Archiv/Review of World Economics* 134(2): 280-301.
- [100] Richardson, M. (1993); "Endogenous Protection and Trade Diversion." *Journal of International Economics*, 34(3-4): 309-24.
- [101] Richardson, M. (1995); "On the Interpretation of the Kemp/Wan Theorem", *Oxford Economic Papers* 47, 696-703.

- [102] Richter, W.F. and D. Wellisch (1996); "The Provision of Local Public Goods and Factors in the Presence of Firm and Household Mobility," *Journal of Public Economics* 60: 73-93.
- [103] Riezman, R. (1985); "Customs Union and the Core," *Journal of International Economics* 19, 355-365.
- [104] Salop, S. (1979); "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10: 141-156.
- [105] Sampson, G.P. (1996); "Compatibility of Regional and Multilateral Trade Arrangements: Reforming the WTO Process", *American Economic Review* 86, 82-88.
- [106] Slikker, M. and A. van den Nouweland (2000); "Communication Situations with Asymmetric Players." *Mathematical Methods of Operations Research*, 52: 39-56.
- [107] Spence, M. (1976); "Product Selection, Fixed Costs and Monopolistic Competition," *Review of Economic Studies* 43: 217-235.
- [108] Spence, M. (1977); "Entry, Capacity, Investment and Oligopolistic Pricing," *Bell Journal of Economics*, 8: 534-544.
- [109] Staiger, R. (1995); "A Theory of Gradual Trade Liberalization." In Levinsohn, J., A.V. Deardorff and R.M. Stern (eds.) *New Directions in Trade Theory*, University of Michigan Press, Ann Arbor, MI.
- [110] Syropoulos, C. (2001); "Optimum Tariffs and Retaliation Revisited: How Country Size Matters." To appear in *Review of Economic Studies*.
- [111] Tiebout, C.M. (1956); "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64: 416-24.

- [112] Whalley, J. and C. Hamilton (1996); *The Trading System After The Uruguay Round*, Institute for International Economics, Washington, DC.
- [113] Wildasin, D.E. (1988); "Nash Equilibria in Models of Fiscal Competition." *Journal of Public Economics*, March 1988, 35(2): 229-240.
- [114] Wildasin, D.E. (1989); "Interjurisdictional Capital Mobility: Fiscal Externalities and a Corrective Subsidy." *Journal of Urban Economics*, 25: 193-212.
- [115] Wilson, J.D. (1986); "A Theory of Interregional Tax Competition." *Journal of Urban Economics*, 19, pp. 296-315.
- [116] Wilson, J.D. (1987); "Trade, Capital Mobility, and Tax Competition." *Journal of Political Economy*, (August) 95(4): 835-856.
- [117] Wilson, J.D. (1991); "Tax Competition with Interregional Differences in Factor Endowments." *Regional Science and Urban Economics*, 1991, 423-451.
- [118] Wilson, J.D. (1986); "Theories of Tax Competition." *National Tax Journal*, (June) 52(2) pp. 269-305.
- [119] Wooders, M.H. (1980); "The Tiebout Hypothesis; Near Optimality in Local Public Good Economies." *Econometrica* 48: 1467-85.
- [120] Wooders, M.H. (1985); "A Tiebout Theorem," University of Minnesota IMA, Preprint 128, published in *Mathematical Social Sciences*, 1989, 18: 33-55.
- [121] Wooders, M. (1999) "Multijurisdictional Economies, The Tiebout Hypothesis and Sorting," *Proceedings of the National Academy of Sciences, USA*, 96: 10585-10587 (September).
- [122] Wooders, M., B. Zissimos and A. Dhillon (2001); "Tax Competition Reconsidered." Warwick Economic Research Papers no. 622, December 2001, 52 pages. Available on-line at <http://www.warwick.ac.uk/economics/research/twerps.html>.

- [123] WTO (2000); "Mapping of Regional Trade Agreements" World Trade Organization, Geneva. Available online at <http://www.wto.org>.
- [124] Yi, S.-S. (1996) "Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism is Good" *Journal of International Economics*, 41: 151-175.
- [125] Zodrow, G.R. and P. Mieszkowski (1986); "Pigou, Tiebout, Property Taxation and the Underprovision of Local Public Goods." *Journal of Urban Economics*, 19: 356-370.